Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

DELOS D. WICKENS, Associate Editor
Ohio State University

Consulting Editors

E. JAMES ARCHER, *University of Wisconsin*
JUDSON S. BROWN, *State University of Iowa*
CLETUS J. BURKE, *Indiana University*
PAUL M. FITTS, *Ohio State University*
FREDERICK C. FRICK, *Massachusetts Institute of Technology*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
LLOYD G. HUMPHREYS, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*

CONTENTS

Counting Repeated Light Flashes as a Function of Their Number, Their Rate of Presentation, and Retinal Location Stimulated: D. M. FORSYTH AND A. CHAPANIS 385

A Procedure for Increasing Word Association Originality and Its Transfer Effects: I. MALTZMAN, W. BOGARTZ, AND L. BERGER 392

The Facilitation of Problem Solving by Prior Exposure to Uncommon Responses: I. MALTZMAN, L. O. BROOKS, W. BOGARTZ, AND S. S. SUMMERS 399

Human Delayed-Reward Learning With Different Lengths of Task: C. E. NOBLE AND W. T. ALCOCK 407

Concept Identification as a Function of Completeness and Probability of Information Feedback: L. E. BOURNE, JR., AND R. B. PENDLETON 413

Transfer Effects on a Rotary Pursuit Task as a Function of First-Task Difficulty: D. S. LORDAHL AND E. J. ARCHER 421

Effect of Distribution of Practice on a Component Skill of Rotary Pursuit Tracking: E. J. ARCHER 427

Effect of Stimuli Time Relations During Preconditioning Training Upon the Magnitude of Sensory Preconditioning: D. R. HOFFELD, R. F. THOMPSON, AND W. J. BROGDEN 437

A Comparison of Correction and Modified Correction Procedures on the Acquisition of a 12-Unit Verbal Maze: R. F. THOMPSON 443

Alley Length and Time of Food Deprivation in Instrumental Reward Learning: N. F. BESCH AND W. F. REYNOLDS 448

Correction and Reanalysis: N. H. ANDERSON AND D. A. GRANT 453

Erratum 454

Supplementary Report: Processes Underlying Learning a Single Paired-Associate Item: J. O. COOK 455

Supplementary Report: The Diminishing Interval Conditioning Technique: R. E. MILLER AND J. V. MURPHY 456

American Psychological Association

Vol. 56, No. 5 November 1958

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Promotion Manager
SADIE J. DOYLE, Editorial Assistant

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, 53, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association*.** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1958 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. and R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1958 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 56, NO. 5 NOVEMBER, 1958

COUNTING REPEATED LIGHT FLASHES AS A FUNCTION OF THEIR NUMBER, THEIR RATE OF PRESENTATION, AND RETINAL LOCATION STIMULATED [1]

D. M. FORSYTH AND A. CHAPANIS

Johns Hopkins University

Several recent experiments have been concerned with the variables affecting an *O*'s perception of the numerousness of repeated stimuli. Auditory (**3, 4, 7**), tactual,[2] and visual stimuli[2] (**2, 5, 8, 9**) have been so studied. In these investigations varying numbers of successive stimuli were presented at different rates. The *O* was instructed to count, judge or estimate the number of stimuli presented. In general, *O*s underestimate the number of stimuli presented and this underestimation increases as the rate of stimulus presentation increases.

The present experiment is concerned with the counting of repeated light flashes as a function of their number, their rate of presentation, and retinal location stimulated. The first two of these variables cover a wider range of values than has previously been investigated in any single study. The third variable—retinal location—has not been included in earlier experiments of this sort.

METHOD

Subjects.—Thirty young men, 18 to 26 yr. old, and 6 young women, 19 to 24 yr. old, were *O*s. All were volunteers who were compensated for their services. None of them wore glasses.

Apparatus.—Flashes of light were produced by a Sylvania R1131C glow-modulator tube. On and off triggers generated by two variable-frequency oscillators produced electric square waves which were the inputs to the flash tube. The rise and decay characteristics of the tube are such that for all practical purposes the pulses of light were square waves. A Hewlett-Packard electronic counter provided precise readings of on-off times, and a scaler gate permitted *E* to select the exact number of flashes to be presented on any given trial. At all frequencies "on" time equalled "off" time.

The test field, mounted in the center of a semicylindrical surround, subtended ½ degree of visual angle. Luminance of the test field was approximately 1800 mL.; luminance of the surround was approximately 22 mL. All observations of the stimulus were monocular (right eye). A head rest helped *O* keep his head in the same position on all trials. All retinal displacements are measured from the fovea (0°) in the horizontal plane. They were

[1] This study was done under Contract N5-ori-166, Task Order 1, between the Office of Naval Research and Johns Hopkins University. This is Report No. 166-I-212, Project Designation No. NR 145-089, under that contract. Reproduction in whole or in part is permitted for any purpose of the United States Government. A summary of this study was presented at the 1957 meetings of the American Psychological Association (1).

[2] C. T. White. Personal communication, 1957.

obtained by having *O*s fixate red point-sources of light at appropriate angular displacements (temporally) from the test field in the center of the surround, the position of the head remaining fixed. Thus the light always stimulated the temporal retina. Except for trials at the fovea, the one correct fixation point during any given trial was lighted by *E*.

Procedure.—Six frequencies, 2.5, 5, 10, 15, 22.5, and 30 cps, and six retinal displacements, 0° (fovea), 2.5°, 5°, 10°, 20°, and 40°, were used. Every *O* was given 20 trials (one for each of 1 to 20 flashes of light) at every one of the 36 combinations of frequency and retinal displacement. Thus each *O* made 720 judgments—$20 \times 6 \times 6$.

The procedure was as follows: Each *O* participated in six experimental sessions, one on each of six different days, generally within a two-week period. On any one day *O* received 120 trials, all at one frequency, i.e., the 1 to 20 flashes at each of the six retinal displacements. The 120 combinations of retinal position and number of flashes were presented in a random order but with the restriction that each combination appeared only once in the sequence. The order of the frequencies used on the six days was different for every *O*, but each frequency appeared first in the series for 6 *O*s, second in the series for 6 *O*s, etc. Experimental sessions lasted about 20 min.

Instructions.—The *O*s were instructed to estimate the number of flashes they had seen, but were given no information about the range of numbers to be expected. They were further instructed that they would usually not be able to count the flashes, but were to make the best estimate possible as to how many there had been.

After receiving a ready signal from *E*, *O* initiated each train of light flashes by pushing a button. The instructions emphasized that they were not to do this until they were fully prepared to observe.

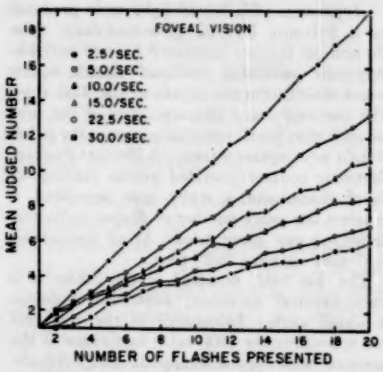

FIG. 1. Mean judged number as a function of number of flashes presented at the fovea. Each point represents the mean of 36 observations, one for each of 36 *O*s. Each curve is for a different rate of presentation.

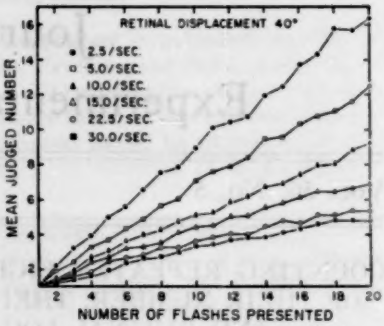

FIG. 2. Mean judged number as a function of number of flashes presented at a retinal displacement 40° temporally from the fovea.

RESULTS

Figure 1 shows a plot of mean judged number vs. number of light flashes presented at the fovea. At 2.5 cps, the mean judged number for 20 flashes is 19.2. The slope of that curve is very nearly 1. This means that at 2.5 cps the average estimates are very accurate. As frequency is increased the slopes of the curves decrease, indicating greater and greater constant errors of estimation. For example, when 20 flashes are presented at 30 cps, the mean judged number is only 6.1.

Figure 2 is similar to Fig. 1, but for a retinal displacement of 40°. A comparison of the two figures shows that the effect of the displacement has been to lower the slopes of the curves at all frequencies. Constant errors of estimation increase with increasing retinal displacement.

Plots similar to those in Fig. 1 and

2 were prepared for retinal displacements of 2.5°, 5°, 10°, and 20°. The method of least squares was used to fit straight lines to the resultant 36 regression lines. The slopes of these straight lines provide a convenient measure of the effects of frequency and retinal displacement on these functions. The 36 slopes appear in Table 1.

Constant errors of estimation.—The slopes in Table 1 are all less than 1. This means that *O*s typically underestimate the number of flashes presented. Of a total of 25,920 judgments, only 446 were overestimates, whereas 21,344 were underestimates. The majority of the overestimates (320) occurred at 2.5 cps. Overestimations decrease rapidly as frequency increases. Of the 4,320 judgments at 30 cps only 4 were overestimates. At 2.5 cps overestimations are relatively independent of the number of flashes presented. At higher frequencies overestimates occur only for small numbers of flashes.

Table 2 gives r^2 and η^2 for all regression lines. The linear component of all regression lines is significant beyond the .001 level.

TABLE 1

SLOPES OF THE REGRESSION LINES RELATING MEAN JUDGED NUMBER TO PRESENTED NUMBER

Retinal Displacement	Frequency in cps					
	2.5	5	10	15	22.5	30
0°	.947	.648	.507	.397	.306	.266
2.5°	.911	.641	.478	.376	.282	.250
5°	.902	.632	.470	.363	.272	.242
10°	.878	.625	.443	.358	.267	.245
20°	.846	.605	.432	.335	.258	.239
40°	.795	.584	.412	.311	.231	.210

Asterisks identify those curvilinear components which are statistically significant. It is plain that even when the curvilinear component is significant, the variance not accounted for by the linear component is due mainly to error and not to curvilinearity. The largest difference between r^2 and η^2 is only .025. The values in Table 2 should be interpreted with some caution because of the heterogeneity of variance in these data (see Fig. 6 and the accompanying discussion).

Effects of frequency.—In Fig. 3, the slopes of the regression lines are plotted against frequency. Freehand curves have been fitted through

TABLE 2

r^2 AND η^2 FOR THE REGRESSION LINES RELATING MEAN JUDGED NUMBER TO PRESENTED NUMBER

Retinal Displacement	Frequency in cps											
	2.5		5		10		15		22.5		30	
	r^2	η^2	r^2	η^2	r^2	η^2	r^2	η^2	r^2	η^2	r^2	η^2
0°	.978	.978	.899	.900	.769	.771	.705	.713	.691	.704*	.692	.716***
2.5°	.966	.967*	.873	.874	.782	.787	.684	.693	.697	.711*	.711	.728***
5°	.959	.961	.866	.868	.798	.801	.686	.696	.725	.739**	.710	.731***
10°	.934	.937*	.855	.856	.778	.784	.704	.711	.687	.705***	.690	.715***
20°	.889	.892	.836	.839	.756	.763	.668	.679	.674	.695***	.664	.682**
40°	.847	.850	.781	.784	.749	.752	.633	.642	.626	.640	.603	.616

* $.05 > P > .01$, for curvilinear component.
** $.01 > P > .001$, for curvilinear component.
*** $P < .001$, for curvilinear component.

only the extreme points to show the general shape of the function. Despite the small range at each frequency, the ordering of the slopes for the different retinal displacements is highly consistent. The most important thing to note from this graph is that the effect of frequency on the ability of an observer to follow a train of light flashes is roughly hyperbolic (log-log transforms of these curves yield very good straight lines).

Effects of retinal location.—Figure 4 shows slopes plotted against retinal displacement. The curves in Fig. 4 were drawn by visual inspection and are subject to a certain amount of arbitrariness. We regret especially not having collected data at a retinal displacement of 30°, but we had not anticipated finding trends like those shown in the figure. Despite this limitation we feel that two important relationships are apparent in these data: (*a*) The curves are two-limbed with a fairly clear-cut separation between the central and peripheral areas of the eye, and (*b*) as frequency increases the breaks in these functions occur farther out in the periphery of the eye.

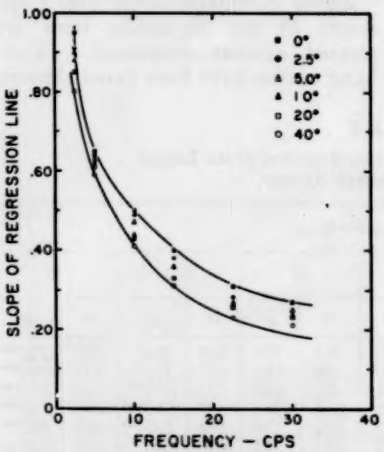

FIG. 3. The slopes in Table 1 are plotted here as a function of frequency. The different symbols identify the 6 retinal displacements.

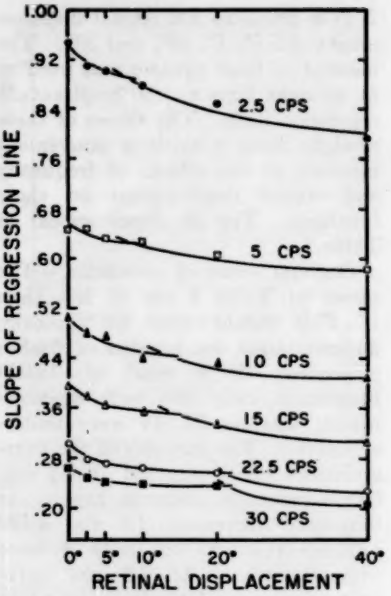

FIG. 4. The slopes in Table 1 are plotted here as a function of retinal displacement. The different curves are for different rates of presentation.

Estimations as a function of total elapsed time.—Cheatham and White (2) and White[3] have used the concept of "subjective rate" or "perceptual rate" in explaining their findings. They believe that this subjective rate has an upper limit of 6–8 pulses per second. They further assert that the subjective rate is determined primarily by some temporal patterning process in the central nervous system and not by the objective stimulus rate. To support this point of view they have plotted their data in terms of the total elapsed time occupied by a train of pulses. In Fig. 5 the data from Fig. 1 are similarly replotted

[3] See Footnote 2.

as a function of total elapsed time. Certain points at the lower frequencies have been omitted. Line *a* represents the slope which would result if the subjective rate were 6 flashes/sec. We are concerned merely with the slope of this line, not with its point of origin.

Our data are in substantial agreement with those of Cheatham and White. When the objective flash rate is 22.5 or 30 flashes/sec. and the trains of pulses occupy less than 300 msec., the over-all slopes of the curves correspond to subjective rates of about 13 flashes/sec. For objective rates of 15, 22.5, and 30 flashes/sec. and for trains of pulses which are more than 300 msec. in duration, the subjective rate is almost exactly 6 flashes/sec. For an objective rate of 10 flashes/sec. the subjective rate is slightly less than 6 flashes/sec. The other two curves are not comparable since the objective rates are both less than 6 flashes/sec.

Plots similar to those of Fig. 5 have been constructed for the other retinal locations. In general, two trends are apparent: (*a*) The very steep slope during the first 300 msec. is gradually reduced until at a retinal

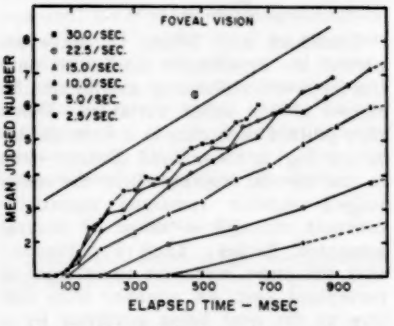

FIG. 5. Mean judged number as a function of total elapsed time from the start of a train of light flashes. Each point represents the mean of 36 estimates.

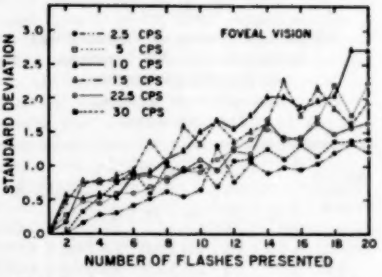

FIG. 6. Each point is a standard deviation of the 36 responses made for each combination of number of pulses and stimulus rate. These data are for foveal vision only.

displacement of 40° the slopes of the curves are nearly constant from beginning to end. (*b*) Although displaced downward in amplitude, the curves have over-all *slopes* almost identical to those in Fig. 5 for durations above 300 msec.

Variability of the estimates.—The variability of the responses increases fairly regularly with an increase in the number of flashes presented. Figure 6 is for foveal vision only, but this general statement holds for all retinal displacements.

In Fig. 6 the frequencies exhibiting the greatest variability are 10 and 15 cps, those exhibiting the least variability are 2.5 and 30 cps, while 5 and 22.5 cps fall roughly in the middle. There is, however, an interaction between frequency and retinal displacement that has an effect on the variability. In order to quantify this effect without becoming involved in complicated plots, the *SD*'s in the data were rank-ordered in two ways. In the first instance we worked from Fig. 6 and from similar plots for the other five retinal displacements. The first step was to rank-order the frequencies at each number of flashes presented, assigning 1 to the largest and 6 to the smallest *SD* at each number. The ranks for each frequency

TABLE 3

MEAN RANK ORDER OF VARIABILITIES AVERAGED OVER ALL NUMBERS OF FLASHES PRESENTED: ANALYZED BY RETINAL DISPLACEMENTS

Retinal Displacement	Frequency in cps						W
	2.5	5	10	15	22.5	30	
0°	5.7	3.5	1.9	2.0	3.4	4.5	.66***
2.5°	5.1	2.4	2.3	1.9	4.4	5.0	.64***
5°	4.6	2.2	2.3	2.1	4.4	5.4	.62***
10°	3.2	2.3	2.5	3.1	4.4	5.6	.46***
20°	2.0	2.4	3.5	2.8	5.0	5.3	.56***
40°	1.7	2.3	3.3	3.1	5.2	5.4	.65***

*** $P < .001$, for Kendall's W.

were then averaged across all numbers of flashes. These are the entries in Table 3. The entry W is Kendall's coefficient of concordance. Line 1 of Table 3, therefore, summarizes the data presented in Fig. 6.

Since the W's in Table 3 are all moderately high and statistically significant, this indicates that the variabilities for the different frequencies tend to be consistent at any retinal location. Table 3 also shows the interaction between frequency and retinal location. This is reflected in the changes of the mean ranks within columns. At the fovea 10 and 15 cps have the greatest, and similar variabilities, while 2.5 and 30 cps have the least. As retinal displacement increases, the pattern shifts until at 40° there is a fairly regular decrease in variability from 2.5 through 30 cps.

The second method of analyzing the variability data also started with six sets of curves in which SD's were plotted against number of flashes presented. This time, however, there was one set for each frequency. Each set in turn consisted of six curves, one for each retinal displacement. Within each set of curves we then rank-ordered retinal displacements, assigning 1 to the largest and 6 to the smallest SD for each presented number. The ranks for each retinal displacement were then averaged across all numbers of flashes. These are the entries in Table 4. Whereas Table 3 shows the mean rank for each frequency at any given retinal location, Table 4 shows the mean rank for each retinal location at any given frequency.

TABLE 4

MEAN RANK ORDER OF VARIABILITIES AVERAGED OVER ALL NUMBERS OF FLASHES PRESENTED: ANALYZED BY FREQUENCIES

Frequency in cps	Retinal Displacement						W
	0°	2.5°	5°	10°	20°	40°	
2.5	5.8	4.7	4.2	3.3	1.8	1.2	.89***
5	5.5	3.8	4.1	3.3	2.6	1.6	.50***
10	2.6	3.6	4.2	3.9	3.4	3.2	.09
15	3.0	3.1	3.6	4.4	3.6	3.3	.07
22.5	2.1	3.6	5.0	3.4	3.5	3.3	.25***
30	2.8	4.0	4.7	3.7	2.8	3.0	.17**

** $.01 > P > .001$, for Kendall's W.
*** $P < .001$, for Kendall's W.

There is not as much consistency in Table 4 as in Table 3. This says in effect that the variability in the estimates is affected more by frequency than by retinal location.

DISCUSSION

Cheatham and White (2) have reported an experiment similar to ours but for foveal vision only and for smaller ranges of the other variables. When they plotted their data in a form similar to our Fig. 1, they found distinct steps in the curves, regions where the mean judged number remained essentially constant through a range of several presented flashes. They hypothesized that the steps represent some type of perceptual unit, an advance from one step to the next being governed by a probability-type function where the addition of one more light flash increases the probability that one more subjective flash will be perceived.

In this type of situation where the range of responses is much smaller than the range of stimuli, if an individual *O* is consistent in correlating responses with stimuli, it is inevitable that the curve showing responses as a function of stimuli will contain steps. We found this for each of our *O*s as did the workers at Mount Holyoke (**5**) when they plotted data for individual *O*s. The precise points at which steps occur differ for the various *O*s. Thus, in averaging one observation for each of 36 *O*s at each point we obtain curves which are regular. The curves plotted by Cheatham and White, however, are for 10 observations made by each of 5 *O*s, and so confound the variance within *O*s with the variance between *O*s. In this process the differences between *O*s contribute much less weight to the mean curves. We believe that this may explain why they found step functions and we did not. An additional consideration is that Cheatham and White used highly trained *O*s whereas ours were relatively naive. One might expect that highly trained *O*s would exhibit a considerable amount of internal stability.

The double-limbed curves which result when our data are plotted as a function of retinal displacement were completely unexpected since such functions are normally obtained only when luminosity is one of the variables. The most nearly relevant study we could find is an experiment by Ross (**6**) who measured CFF at retinal displacements from 0° to 45°. He found no evidence of a duplex function, although, of course, his experiment is not completely analogous to the one reported here. We are convinced that our duplex functions are genuine, but at the present time we have no good explanation for them.

Summary

From 1 to 20 flashes of light were presented to *O*s at six frequencies (2.5 to 30 cps) and six retinal displacements (0° to 40°). For the pooled data of 36 *O*s, mean judged number is generally a linear function of the number of flashes presented. The slopes of these straight lines are hyperbolic functions of frequency for all retinal displacements. The slopes of these straight lines plotted against retinal displacement yield discontinuous, two-limbed curves in which the position of the break is a function of frequency. At all frequencies the variability of the responses increases as number of flashes increases. The variability in the estimates is affected more by stimulus frequency than by retinal location.

References

1. Chapanis, A., & Forsyth, D. M. Estimation of light pulses as a function of retinal location. *Amer. Psychologist*, 1957, **12**, 441.
2. Cheatham, P. G., & White, C. T. Temporal numerosity: I. Perceived number as a function of flash number and rate. *J. exp. Psychol.*, 1952, **44**, 447–451.
3. Cheatham, P. G., & White, C. T. Temporal numerosity: III. Auditory perception of number. *J. exp. Psychol.*, 1954, **47**, 425–428.
4. Garner, W. R. The accuracy of counting repeated short tones. *J. exp. Psychol.*, 1951, **41**, 310–316.
5. Reese, E. P., Reese, T. W., Volkmann, J., & Corbin, H. H. Psychophysical research, summary report 1946–1952. Office of Naval Research, *Tech. Rep. SPECDEVCEN 131-1-5*, 1953.
6. Ross, R. T. The fusion frequency in different areas of the visual field: II. The regional gradient of fusion frequency. *J. gen. Psychol.*, 1936, **15**, 161–170.
7. Taubman, R. E. Studies in judged number: I. The judgment of auditory number. *J. gen. Psychol.*, 1950, **43**, 167–194.
8. Taubman, R. E. Studies in judged number: II. The judgment of visual number. *J. gen. Psychol.*, 1950, **43**, 195–219.
9. White, C. T., Cheatham, P. G., & Armington, J. C. Temporal numerosity: II. Evidence for central factors influencing perceived number. *J. exp. Psychol.*, 1953, **46**, 283–287.

(Received January 18, 1958)

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

A PROCEDURE FOR INCREASING WORD ASSOCIATION ORIGINALITY AND ITS TRANSFER EFFECTS[1]

IRVING MALTZMAN, WILLIAM BOGARTZ, AND LOUIS BREGER

University of California, Los Angeles

Studies by Maltzman and his associates (5) have been concerned with examining the relationship between uncommon verbal responses and problem solving and with attempts to facilitate problem solving by evoking uncommon responses. The present study, in contrast, is concerned directly with the facilitation of uncommon verbal responses.

One of the group of originality tests developed by Guilford and his coworkers (2, 7), the Quick Responses Test, is a free association test in which the frequency of occurrence of a response to a given stimulus word is taken as its measure of uncommonness. In terms of S-R theory the response that occurs to a given stimulus in the free association situation is the one which is dominant in the response hierarchy elicitable by that stimulus. It has the greatest amount of excitatory strength at the moment of stimulus presentation. Responses that are uncommon or original must be lower in the response hierarchy, possess a weaker excitatory potential. If, therefore, a situation could be arranged in which *S* is induced to give responses low in his hierarchy, there would be an increase in the originality of his responses. Training of this nature might then produce a disposition to give uncommon responses in other situations. The procedure employed in this study in an attempt to produce the desired effect is the simple one of repeatedly presenting the same stimulus words with instructions to give a different response each time. This is only one of several different procedures that may be used to induce *Ss* to increase the uncommonness of their responses, but it is well suited for use in the free association situation.

Two further experimental variables were introduced in an attempt to facilitate the effect of the primary experimental variable of repeated presentation of the stimulus words and evocation of different responses. These were verbal reinforcement of uncommon responses on a partial reinforcement schedule, and instructions to be original.

Following the free association test, *Ss* were given Guilford's Unusual Uses test of originality[2] (2, 7) in order to estimate the generality of the effects of training and instructions to give uncommon responses. An alternate form of the test was devised and employed with approximately one-third of the *Ss*, but was discarded because the response distributions obtained appeared quite different from those obtained from the Guilford form.

METHOD

Subjects.—The *Ss* were 120 students, 60 men and 60 women, drawn from introductory psychology classes.

[1] This research was supported in part by the United States Air Force under Contract No. AF 41(657)-52, monitored by the Operator Laboratory Field Extension No. 1, Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

[2] We are indebted to Paul R. Christensen for kindly providing us with a copy of this test.

Stimulus materials.—The stimulus words employed in the free association situation were selected from the norms obtained by Wilson (6). He presented words culled from previous word association lists to 300 *Ss*. Presentation of the words was conducted on an individual basis, and each *S* responded only once to each of the stimulus words. Oral presentation of the words was employed, as was the case in this study. Fifty words with the smallest number of different responses associated to them were selected for the training and test lists. The words are presented in Table 1.

The Unusual Uses test asks *S* to consider in order six different common objects—an automobile tire, a key, a safety pin, a watch, a button, and eyeglasses. A common use for each object is described in the test, and *S* is asked to give additional different uses for each of them. Ten minutes is allowed for completion of the test.

Procedure.—The *Ss* were treated in identical fashion on the initial presentation of the free association training list, and received the usual instructions to respond as quickly as possible with the first word that came to mind. After their first completion of the training list the two experimental groups were instructed to again respond as quickly as possible to each word, but to give a different response from the one used before. The training list was repeated five times in this fashion. The two experimental groups receiving repeated presentations of the training list differed only with respect to the administration of verbal reinforcement for responses that were judged to be uncommon. These judgments were based on the analysis of data obtained from an initial exploratory experiment. For the reinforcement group approximately every fifth uncommon response to the training words was called "good" by *E*.

The experimental groups received the test list of new words after the last repetition of the training list while the control group received the test list after their first completion of the training list. Prior to the presentation of the test list half the *Ss* in each of the three groups received instructions to be as original as possible on the new list of words while the remaining *Ss* received the usual free association instructions.

Following completion of the free association test list *Ss* were given the Unusual Uses test. Approximately half the *Ss* within each of the six free association subgroups were given instructions to be as original as possible and half were not.

TABLE 1

STIMULUS WORDS EMPLOYED IN FREE ASSOCIATION TRAINING AND TEST LISTS

Training List		Test List	
1. pencil	13. stork	1. rich	13. stove
2. shut	14. swift	2. needle	14. high
3. dark	15. tobacco	3. ocean	15. voyage
4. scissors	16. blossom	4. slow	16. girl
5. white	17. yellow	5. carpet	17. cow
6. black	18. bake	6. ink	18. fast
7. sour	19. hammer	7. thirsty	19. red
8. blue	20. rough	8. man	20. short
9. boy	21. loud	9. window	21. woman
10. butter	22. frosty	10. knife	22. lamp
11. table	23. brown	11. village	23. hot
12. bitter	24. sweet	12. city	24. frog
	25. mutton		25. whiskey

RESULTS

The frequency with which the different responses occurred to each stimulus word was determined for the training and the test lists. The *Ss*' responses were then scored on the basis of the obtained frequencies, and each *S* was assigned an originality score which was the mean frequency of his 25 responses. A low score therefore represents high originality.

It was assumed that the training administered the experimental groups, i.e., repeated presentation of the initial list with instructions to give different responses, would increase the uncommonness of the responses to the training words. Analysis of the response frequencies on successive presentations of the training list indicated that this condition was obtained. On the average, over 75% of the responses on the fifth repetition were unique, i.e., were given by only one *S*. This represents more than a threefold increase over the initial presentation of the training list. Almost identical results were obtained for the two experimental groups, indicating that the verbal reinforcement employed did not markedly influence the uncommonness of the responses during training.

A factorial analysis of variance of performance on the initial presentation of the training list showed that the various groups did not differ

TABLE 2

UNADJUSTED AND ADJUSTED MEAN FREE ASSOCIATION TEST ORIGINALITY SCORES

Condition	Training		Reinforced Training		Control	
	N	Mean	*N*	Mean	*N*	Mean
Unadjusted						
No instructions	20	20.94	20	20.49	20	23.98
Instructions	20	4.93	20	4.41	20	7.61
Adjusted						
No instructions	20	21.37	20	20.34	20	23.80

significantly prior to introduction of the experimental treatments. Since performance on the initial presentation of the training list and the test list was significantly correlated for *Ss* in the three nonoriginality instruction subgroups ($r = .54$), an analysis of covariance was conducted on these groups. Analysis of variance was employed in a separate evaluation of the three originality instruction subgroups because their test performance was not significantly correlated with performance on the initial presentation of the training list, and their variances were appreciably smaller than the nonoriginality instruction groups.

Table 2 shows the adjusted and unadjusted mean originality scores obtained by the various groups on the test list.

Results of an analysis of covariance (**3,** Type I) of the scores obtained by the experimental and control groups that did not receive originality instructions are shown in Table 3. It is evident that the main effect of training is significant. Also significant are the item effects and the treatment by item interaction, indicating that the stimulus words differed in their tendency to evoke uncommon responses, and that this tendency differed under the control and experimental conditions. Results of *t* tests show that the verbally reinforced experimental training group gave significantly more original responses than the control group ($P = <.02$). The difference between the experimental training group that did not receive verbal reinforcement and the control group is in the same direction, but does not quite reach an acceptable level of confidence

TABLE 3

ANALYSIS OF COVARIANCE OF NO-ORIGINALITY INSTRUCTION GROUPS ON THE FREE ASSOCIATION TEST LIST

Source	*df*	*MS*	*F*	*P*
Treatments (T) + Error (b)				
Error (b)	56	499.51		
Treatments	2	1604.95	3.21	.05
Items (I) + Error (w)				
Error (w)	1367	187.82		
Items	24	5482.95	29.19	.0001
T × I + Error (w)				
Error (w)	1367	187.82		
T × I	48	342.95	1.83	.01

($.05 < P < .10$). The two experimental groups do not differ significantly.

Results of an analysis of variance (**3**, Type I) of the test list scores obtained by the experimental and control groups receiving originality instructions are shown in Table 4. Significant training and item effects and a treatment by item interaction were again obtained. The results of *t* tests show that the nonreinforced experimental group gave significantly more original responses than the control group ($P = <.01$). The reinforced experimental group again is significantly superior to the control group ($P = <.01$) while the two experimental groups do not differ significantly.

It is evident from an inspection of the means that originality instructions produced a marked increase in the uncommonness of response and greatly reduced the variance. A *t* test employing half the appropriate degrees of freedom was used to evaluate the difference between the two instruction conditions (**3**). The difference between the obtained means is highly significant ($P = <.0001$).

Originality scores for performance on the Unusual Uses test were obtained by a procedure similar to the one employed for the free association test. The frequency with which different relevant uses occurred for each of the six objects mentioned in the test was first determined. The *Ss*' uses were then rescored on the basis of the obtained frequencies. Each *S*'s originality score was taken as the sum of the weighted uses divided by his total number of uses offered. The total number of uses was also taken as a measure of fluency.

The fluency and originality measures derived from the Unusual Uses test in this manner did not correlate significantly with the originality scores obtained from the initial presentation of the free association training list. These data were therefore evaluated by $3 \times 2 \times 2$ factorial analyses of variance rather than covariance. A summary of the analysis of variance of the fluency scores is shown in Table 5. The only significant effect obtained is the interaction of free association training with originality instructions given prior to the free association test. Inspection of the mean fluency scores in Table 6 suggests that originality instructions given prior to the association test in conjunction with the experimental training are related to a decrease in

TABLE 4

ANALYSIS OF VARIANCE OF ORIGINALITY INSTRUCTION GROUPS ON THE FREE ASSOCIATION TEST LIST

Source	*df*	*MS*	*F*
Between *Ss*	59		
Treatments	2	1474.40	5.89**
Error (b)	57	250.52	
Within *Ss*	1438		
Items	24	572.71	7.16**
T × I	48	122.18	1.53*
Error (w)	1366	80.00	
Total	1497		

* $P = .05$.
** $P = .005$.

TABLE 5

ANALYSIS OF VARIANCE OF UNUSUAL USES FLUENCY

Source	*df*	*MS*	*F*
Treatments (T)	2	38.34	
Instructions 1 (I_1)	1	59.03	
Instructions 2 (I_2)	1	19.63	
T × I_1	2	111.85	4.20*
T × I_2	2	24.32	.91
I_1 × I_2	1	60.57	2.27
T × I_1 × I_2	2	72.99	2.74
Within cells	81	26.66	
Total	92		

* $P < .05$.

TABLE 6

MEAN UNUSUAL USES FLUENCY

Instructions	Training		Reinforced Training		Control	
	N	Mean	*N*	Mean	*N*	Mean
None	6	21.67	7	17.43	8	13.88
Free association	7	11.29	8	12.75	7	18.00
Unusual uses	8	15.13	8	12.75	9	16.11
Free association and unusual uses	8	13.63	9	13.44	8	16.88

the fluency of unusual uses. There is also apparent a tendency for fluency to increase as a function of originality training when originality instructions are absent. These results, however, must be viewed with caution since the cell frequencies in the analyses of variance are disproportionate.

Analysis of variance of the originality scores derived from the Unusual Uses test in the manner described above showed that none of the treatments or their interactions is significant. It was reasoned, however, that the scoring method used may not give adequate weight to the uncommon responses. An alternative scoring method was therefore investigated. However, selection of a new scoring procedure in this manner capitalizes on chance variations in the scores, and the results obtained from the reanalysis must also be viewed with caution. The new originality scores were obtained by determining the number of uses given by each *S* which had a frequency of 10% or less. The mean number of uncommon responses scored in this fashion are shown for each group in Table 7.

TABLE 7

MEAN NUMBER OF ORIGINAL USES

Instructions	Training		Reinforced Training		Control	
	N	Mean	*N*	Mean	*N*	Mean
None	6	10.50	7	8.29	8	4.13
Free association	7	5.00	8	6.25	7	9.14
Unusual uses	8	7.13	8	5.50	9	8.89
Free association and unusual uses	8	6.38	9	6.56	8	7.75

TABLE 8

ANALYSIS OF VARIANCE OF ORIGINAL USES

Source	*df*	*MS*	*F*
Treatments (T)	2	6.18	.35
Instructions_1	1	4.77	.27
Instructions_2	1	.01	.00
$T \times I_1$	2	39.63	2.25
$T \times I_2$	2	21.59	1.23
$I_1 \times I_2$	1	.22	.01
$T \times I_1 \times I_2$	2	67.72	3.85*
Within cells	81	17.60	
Total	92		

* $P < .05$.

A summary of the analysis of variance of these originality scores is shown in Table 8. Only the triple-order interaction was found to be significant. Simple analyses of variance were conducted to determine the locus of the interaction. The results of these tests show that, in the absence of originality instructions, training in the free association situation produced a significant increase in originality on the Unusual Uses test ($P = .05$). As indicated in Table 7, the mean originality scores in the first column are significantly different. The means in each of the remaining columns do not differ significantly within the column. The means in the first and in the last row differ significantly within the respective rows ($P = .05$). Originality instructions in conjunction with originality training without verbal reinforcement appear to have produced a significant decrease in originality on the Unusual Uses test. A similar trend may be seen in the originality training condition with verbal reinforcement, but the decrement is not significant. Originality instructions produced the opposite effect in the control condition. In the absence of originality training in the free association situation, instructions to be original appear to have facilitated the production of original uses.

These interpretations are highly tentative since they are based upon a reanalysis of data, as previously indicated. That these differences could be due to chance variations is even more likely in view of their appearance only as a triple-order interaction. This is especially the case where the cell frequencies are disproportional as they are here.

DISCUSSION

The results obtained clearly suggest that the originality of word associations

to new words can be facilitated to a significant degree. Repeated evocation of different responses to the same training words as well as instructions to be original produced a significant increase in the uncommonness of the responses evoked on the word association test list. The latter list was composed of words not encountered on the training list. Further experimentation is necessary in order to determine whether other training methods besides evocation of different responses to the same stimulus words may facilitate the production of original responses to new words. Verbal reinforcement of uncommon associations did not produce a significant increase in originality beyond that of the training effect. However, it is possible that a greater number of reinforcements and/or a different reinforcement schedule would produce a more pronounced effect than the one obtained.

Instructions greatly facilitated the uncommonness of free associations on the test list for all groups, but the training effect appeared to be greater under the influence of instructions to be original than in the absence of such instructions. Results of this kind are analogous to the effects of instructions in mental set studies, and could be interpreted in a similar fashion (**4**).

Effects of instructions and originality training with word associations on subsequent performance on the Unusual Uses test were ambiguous. Training in conjunction with instructions on the free association test appeared to inhibit the fluency of uses. A similar effect appeared in the second scoring of the originality of uses. Since these effects appeared only as interactions, and the cell frequencies were disproportional, further research in which instructions are manipulated with a variety of test materials is necessary before any general conclusions can be reached about the inhibiting or facilitating effects of originality instruction (**1**).

Evidence of the facilitating effect of training with word associations on the originality of uses also is highly ambiguous, and must be confirmed by further experimentation. If such confirmation is forthcoming, it would imply that the kinds of mediating and associative linkages among verbal responses are considerably more complex than those generally considered in current interpretations of the higher mental processes. The finding in the first phase of this experiment that training on one list of words may facilitate the uncommonness of responses to another list itself indicates that nonspecific transfer may be more readily produced than has been commonly assumed.

Summary

Two experimental groups and a control group were initially presented with a word list and free associations were evoked. The control group then received a new list of stimulus words which served as the test list. The two experimental groups received five additional presentations of the first list with instructions to give a different response on each presentation. One experimental group received intermittent verbal reinforcements for uncommon responses while the other did not. Both groups then received the test list. Half the *Ss* within each of the three groups received originality instructions prior to the test list.

The two experimental groups did not differ significantly in the originality of their associations on the test list. They were significantly more original than the control group under instructions to be original. The verbal reinforcement group was significantly superior to the control group under no originality instructions, and the difference between the control and nonreinforcement groups approached significance. Instructions produced a highly significant increase in originality.

The Unusual Uses test was administered following the free association test. Approximately half the *Ss* receiving this test were given originality instructions prior to its administration. Free association training and originality instructions appeared to produce significant interactions in the number of unusual uses and originality when the latter was rescored. These effects, however, require further confirmation before they can be interpreted unambiguously.

References

1. Christensen, P. E., Guilford, J. P., & Wilson, R. C. Relation of creative responses to working time and instructions. *J. exp. Psychol.*, 1957, **53**, 82–88.

2. Guilford, J. P. Creativity. *Amer. Psychologist*, 1950, **5**, 444–454.
3. Lindquist, E. E. *Design and analysis of experiments in psychology and education*. New York: Houghton Mifflin, 1953.
4. Maltzman, I. Thinking: From a behavioristic point of view. *Psychol. Rev.*, 1955, **62**, 275–286.
5. Maltzman, I., Brooks, L. O., Bogartz, W., & Summers, S. S. The facilitation of problem solving by prior exposure to uncommon responses. *J. exp. Psychol.*, 1958, **56**, 399–406.
6. Wilson, D. P. An extension and evaluation of association word lists. Unpublished doctor's dissertation, Univer. Southern California, 1942.
7. Wilson, R. C., Guilford, J. P., & Christensen, P. R. The measurement of individual differences in originality. *Psychol. Bull.*, 1953, **50**, 362–370.

(Received January 2, 1958)

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

THE FACILITATION OF PROBLEM SOLVING BY PRIOR EXPOSURE TO UNCOMMON RESPONSES [1]

IRVING MALTZMAN, LLOYD O. BROOKS, WILLIAM BOGARTZ, AND STANLEY S. SUMMERS

University of California, Los Angeles

Guilford and his co-workers (**4, 13**) have suggested that the frequency of uncommon but relevant responses may be employed as one of several objective measures of originality. It has also been suggested that originality measured in this fashion may be related to the ability to solve "insight" problems (**4**). The principal purpose of the first experiment reported here was to determine whether such a relationship could be obtained.

According to S-R interpretations of problem solving (**2, 7, 9**) an *S* enters a problem situation with a hierarchy, or hierarchies, of responses which differ in their tendency to occur in that situation. A problem, by definition, cannot be solved by the dominant responses, the responses with the greatest probability of occurrence. Original problem-solving behavior therefore occurs through the arousal of responses relatively low in *S*'s response hierarchy.

If this analysis is essentially correct, then the problem of increasing originality is one of increasing the disposition to produce uncommon but relevant responses. A number of training procedures may be employed to evoke uncommon responses and possibly facilitate originality as a consequence. One procedure is to systematically vary the situation in order to induce *S* to emit progressively more uncommon responses. This technique has been used by Maltzman, Bogartz, and Breger (**8**) with a word association task.

A second procedure is to elicit directly the uncommon responses by stimulus materials presented to *S*. The effects of this latter technique were the primary concern of the last three experiments reported here. The specific hypothesis investigated by these experiments is that, if relatively uncommon responses are elicited, even though they do not enter into the solution of a subsequent problem, their elicitation would facilitate the occurrence of a problem solution. Maier's (**6**) two-string problem was chosen as the problem task in all the experiments, because it is a problem commonly assumed to involve original thinking (**3, 6**), and it lends itself readily to the training technique employed for eliciting uncommon responses.

EXPERIMENT I

Method

A procedure similar to Guilford's Unusual Uses test (**4, 13**) was employed in an attempt to study the relationship between originality and problem solving. Different uses for designated objects are suggested by *S* in this test. The frequency with which the different uses occur can then be determined, and each use offered by *S* can be weighted accordingly. Thurstone's S-L test of verbal fluency (**5, 12**) was administered and scored in terms of uncommonness of response in an attempt to obtain a

[1] This research was supported in part by the United States Air Force under Contract No. AF 41 (657)-52, monitored by the Operator Laboratory Field Extension No. 1, Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

relatively unambiguous and efficient measure of originality.

The susceptibility of the uses test to instructions to be original was also investigated along with the effects of variations in a reference object of the associated uses. These experimental variations were studied by employing a 2 × 2 factorial design in which half of the experimental groups were given originality instructions while half were not. Half of the *Ss* within each of these two groups were shown a large screwdriver while half were shown a small screwdriver as a referent for their associations.

Subjects.—The *Ss* were 181 students, 114 women, and 67 men, drawn from introductory psychology classes. There were 35 *Ss* in Group I and I-0, 39 in Group II, 38 in Group II-0, and 17 in Group III and in Group IV. It was initially planned to run 40 *Ss* in each of the four main experimental groups, and 20 in each of the two control groups. However, the experiment was terminated after only 181 *Ss* were obtained because continuation of the experiment would have coincided with discussion of problem solving in *Ss*' classes which might have differentially affected performance in the experiment.

Procedure.—The six main groups differed in their treatment prior to performing on the two-string problem. All *Ss* initially received the S-L test of verbal fluency. Instructions for the test were to write in 3 min. as many different words as possible that began with the letter "S" and ended with the letter "L." Experimental *Ss* were then given the uses test. A screwdriver, a block of balsa wood, and a length of string suspended from the ceiling were presented individually to each *S*. The *Ss* were instructed to list on a sheet of paper as many different uses as they could think of for each object in 5 min. The order of presentation of the objects to the different *Ss* was varied according to a prearranged random pattern.

For Group I, a large screwdriver 11.5 in. long was presented as their reference for the associated uses and as the implement available in the two-string problem. A small screwdriver 7 in. long was employed in its place for Group II. Groups I-0 and II-0 received the large and small screwdriver, respectively, with additional instructions to be as original as possible in the uses suggested for the three stimulus objects. Groups III and IV received the two-string problem immediately after the S-L test, thus serving as controls for any possible effects that performance on the uses test might have on solution of the two-string problem. Group III received the large screwdriver and Group IV the small screwdriver. All groups received the uses test after completion of the two-string problem. The experimental groups received their respective preproblem instructions while the two control groups received the same instructions as Groups I and II.

The two-string problem was presented in the following fashion. Two strings of equal length were suspended from the ceiling reaching to 45.5 in. from the floor in diagonal corners of the room. A distance of approximately 14 ft. separated the strings. The appropriate screwdriver and the block of balsa wood, approximately 1.75 × 1.5 × 1 in., were placed on a table along one wall of the room. The *Ss* were instructed that only the objects on the table could be employed in the problem.

Solution of the problem required tying the screwdriver to the end of one of the strings, swinging it, extending the other string to its maximum length, and attempting to catch the first string on its upswing. The *Ss* were allowed to work on the problem for 10 min. and the measure of problem-solving behavior was the amount of time required to solve the problem. Solution times were recorded to the nearest second by means of a stop watch. The *Ss* who failed to solve the problem within the allotted time were given a score of 600.

Results

In studying the uses proposed for the string and for the balsa wood, it was noted that *Ss* were apparently inconsistent in their interpretation of the instructions. Some *Ss* restricted their uses to the specific object presented to them. Others did not, and proposed uses that would require large amounts of balsa wood or rope, making it impossible to evaluate the originality of the uses offered by different *Ss* for these two objects. Since this ambiguity did not appear in the uses offered for the screwdrivers, only the results obtained with these stimuli were analyzed for originality.

The basis for scoring originality was the classification of each use as either a member of the class of most common functions of the object or a nonmember of that class. For the screwdriver, this procedure classified each association as either a tool use or a

nontool use. Analysis of the frequency of occurrence of the tool and nontool classes of uses showed a marked difference, substantiating their designation as common and uncommon associations. The mean number of tool uses was 8.5 ($SD = 2.8$) while the mean number of nontool uses was 1.5 ($SD = 1.7$).

According to S-R interpretations of problem solving, the uses produced by each *S* represent a hierarchy of responses differing in excitatory potential. Uncommon nontool uses which have a relatively weak excitatory potential should therefore occur relatively late in the list of uses offered by *S*. This notion is supported by an analysis of the frequency of occurrence of the nontool uses in the first and second halves of each *S*'s list of uses. The mean number of uses given in the first half was 1.4 ($SD = .73$) while the mean number of uses for the second half was 1.9 ($SD = 1.10$). A simple sign test of these data yielded a z of 6.2, $P = .001$. Christensen, Guilford, and Wilson (1) have recently reported comparable results, finding that the frequency of uncommon responses increases with response order in the Unusual Uses test.

An originality score derived from the screwdriver associations was obtained for each *S* by scoring 1 if it was a tool use and 2 if it was a nontool use. These scores were then summed and divided by the total number of screwdriver uses. For the S-L test, the frequency with which each word occurred was first determined. It was then weighted by its absolute frequency and *S*'s originality score was the sum of these frequencies divided by his total number of words.

Since the distribution of solution times for the two-string problem appeared highly skewed, the reciprocals of these scores were obtained. The distribution of solution speeds, however, also appeared skewed. Rank-order statistics were therefore used in all of the analyses reported.

Rank-order analyses of variance indicated that the experimental treatments of originality instruction and size of screwdriver did not produce a significant effect upon problem-solving speed or the originality measures derived from the uses test. An H of 3.29, $df = 5$, was obtained for problem-solving speed, and 4.31, $df = 3$ for originality.

In agreement with other studies involving the two-string problem and similar "insight" problems (11), men obtained significantly faster solutions than women ($z = 5.08$, $P = .001$). Men obtained significantly higher originality scores than women on the preproblem uses test ($z = 2.52$, $P = .01$), but not on the postproblem uses test. The U test was employed in these evaluations of the data.

The rank-order correlations, rhos, obtained between problem-solving speed and measures of originality derived from the uses test are shown in Table 1. They indicate that a slight relationship between these measures does hold for the men. A rho of .61 ($P < .01$) was obtained between pre- and postproblem origi-

TABLE 1

RANK-ORDER CORRELATIONS BETWEEN ORIGINALITY SCORES AND PROBLEM-SOLVING SPEED IN EXP. I

Sex	Preproblem Originality and Solution Speed		Postproblem Originality and Solution Speed	
	N	Rho	*N*	Rho
Men	55	.23*	67	.27*
Women	92	.02	114	.08

* $P = .05$.

TABLE 2

USES PRESENTED IN THE CHECK LISTS FOR EXP. II–IV

Cond.	Screwdriver Items	String Items	Balsa Wood Items
		Exp. II	
X	paper cutter ice chopper paint mixer measuring rod paperweight	hang decorations clothesline test strength demonstrate knots measure distances	fishing float dart game pin cushion paperweight child's toy
		Exp. III	
X_1	table leg instead of paperweight	same as Exp. II	same as Exp. II
		Exp. IV	
X_2	table leg instead of paperweight	dog leash instead of hang decorations	table leg prop instead of paperweight
C_1	unscrew screws screw in a bolt fix a toaster fix a light switch fix a car motor	tie a box tie knots tie a package tie things together in place of shoelace	carve a figure whittle carve a model plane carve a doll's shoe carve a model boat

nality scores for women. A rho of .34 ($P = .05$), was obtained between pre- and postproblem originality scores for men. Originality scores obtained from the S-L test do not correlate significantly with either problem-solving speed or uses originality.

EXPERIMENT II

The purpose of this and the subsequent experiments conducted was to study the effects of a training method designed to facilitate solution of the two-string problem.

Method

Subjects.—The *Ss* were 40 college students drawn from introductory psychology classes. There were 8 women and 13 men in the experimental group, and 10 women and 9 men in the control group. The *Ss* were alternately assigned to the two groups in the order of their appearance for the experiment. One *S* was inadvertently assigned to the experimental group instead of the control group.

Procedure.—Both groups first received the S-L test as a warm-up task. Control *Ss* were then given the two-string problem. Following the S-L test, the experimental *Ss* were given three lists of uses, one for each of the three objects subsequently found in the problem situation—a screwdriver, a block of balsa wood, and a length of string. There were five different uses in each of the lists. Each use in a given list was paired with every other, giving a total of 10 paired items in a list. Arrangement of the items in a list was based on the scheme suggested by Ross (10). The experimental *Ss* were instructed to check which member of each pair they would be more likely to use. The uses appearing in each of the three lists were selected as being relatively uncommon on the basis of a preliminary inspection of the data collected in Exp. I. The items appearing in the lists are presented in Table 2. The order of presentation of the lists was varied according to a prearranged random order. There was no indication to the experimental group that the two-string problem was to follow the check lists.

The two-string problem was presented in the same manner as in Exp. I. The small screwdriver was employed in this and in the subsequent experiments.

Results

Table 3 shows the solution times obtained by the experimental and control *Ss*. Since the distributions of solution times were highly skewed, they were evaluated by the Mann-Whitney *U* test. Experimental men obtained significantly faster solution times than control men ($P = .05$), and experimental women were significantly faster than control women ($P = .04$). Control men were significantly faster than control women ($P = .02$).

Experiment III

This experiment was conducted in order to confirm the stability of the results from the previous study and to investigate additional variables and possible uncontrolled factors.

A new condition was added in order to determine whether the difference previously obtained between the experimental and control groups could be due to transfer from the "paperweight" use in the list for the screwdriver. Since the solution of the two-string problem depends upon using the screwdriver as a weight, a possible criticism of Exp. II is that the problem-solving performance of the experimental group was facilitated by some form of complex transfer from this item. The third group therefore differed from the original experimental group in that this item in the screwdriver list of uses was replaced by the item "table leg."

The second group added in this experiment free associated uses to the screwdriver, balsa wood, and string.

TABLE 3

Solution Times for Experimental and Control Groups in Exp. II

Cond.	Men			Women		
	Mean	Mdn.	*SD*	Mean	Mdn.	*SD*
Experimental	104	41	158	155	72	196
Control	194	123	191	509	585	134

TABLE 4

Solution Times for the Four Groups in Exp. III

Cond.	Men			Women		
	Mean	Mdn.	*SD*	Mean	Mdn.	*SD*
X	47	40	37	169	101	173
X_1	122	52	161	143	90	157
FA	122	72	146	338	324	232
C	140	60	176	320	280	203

Method

Subjects.—The *Ss* were 124 students. There were 18 men and 13 women in the free association (FA) group, 15 men and 16 women in the control (C) group, 19 men and 12 women in the original experimental group (X), and 14 men and 16 women in the revised experimental (X_1) group.

Procedure.—Group C and Group X were treated in the same fashion as in Exp. II, and constituted a repetition of that study. Group X_1 differed from the first experimental group only in the replacement of the item "paperweight" by the item "table leg" in the list of uses for the screwdriver. Group FA associated uses to the objects available in the experimental situation in the same fashion as Group II in Exp. I.

Results

Solution times for men and for women in each of the four conditions are shown in Table 4. A rank-order analysis of variance of the men in the four conditions yielded an *H* of 7.97, $df = 3$, $P = .05$. *U* tests were made between individual groups of men. The only significant difference ($P = .02$) obtained was between Group X and Group FA. The combined experimental groups differed from Group C at the .06 level and Group FA at the .01 level.

A rank-order analysis of variance of the women in the four conditions yielded an H of 11.66, $df = 3$, $P = .01$. U tests were made among the individual groups of women. Group X obtained significantly faster solution times than Group C ($P = .05$). Group X_1 obtained significantly faster solution times than Group C ($P = .02$) and Group FA ($P = .02$). When the two experimental groups are combined, they differ significantly from Group C and Group FA at the .01 level. Men were significantly faster ($P = .01$) than women in all conditions except in Group X_1.

Experiment IV

Equivocal results were obtained in Exp. III with respect to the question of transfer from the check list to the problem-solving situation.

A third experiment was therefore conducted. The original control and experimental conditions of Exp. II were again employed. A revised experimental group was added in which several items were changed in order to avoid possible sources of transfer.

A second control group was added to provide a more direct test of the hypothesis that different uncommon uses were responsible for the facilitation of problem solving. This group received a list of relatively similar common uses for each object. As a consequence, this group should not differ significantly from the original control group, and should be significantly slower than the experimental groups in its solution of the two-string problem.

TABLE 5

Solution Times for the Four Groups in Exp. IV

Cond.	Men			Women		
	Mean	Mdn.	*SD*	Mean	Mdn.	*SD*
X	56	40	50	209	118	208
X_2	64	49	41	198	115	191
C	99	51	158	339	346	223
C_1	104	45	161	310	191	225

Method

Subjects.—The *S*s were 144 students. There were 13 men and 23 women in each of the four groups employed. They were assigned to the different groups in the order of their appearance with the added restriction that the proportion of men and women be equal in the different groups.

Procedure.—The control group (C) and the first experimental group (X) received the same treatment as in the previous experiments. The second experimental group (X_2) differed from Group X in terms of the substitution of new items in the check list as shown in Table 2. The new control group (C_1) was given a list of common uses for each of the objects as indicated in Table 2.

Results

Solution times for men and women in each of the four conditions are shown in Table 5. A rank-order analysis of variance of the solutions times obtained by the men in the four conditions yielded an H of 1.08, $df = 3$, which is not significant. A rank-order analysis of variance of the solution times obtained by the women resulted in an H of 8.43, $df = 3$, $P = .05$. Results of U tests among the individual groups of women showed that the two experimental groups did not differ significantly. Group X was significantly faster than Group C ($P = .02$), and approached significance in comparison with Group C_2 ($P = > .05$). Group X_2 was significantly faster than Group C ($P = .02$) and Group C_1 ($P = .04$). The combined experimental groups were significantly superior to Group C ($P = .01$) and Group C_1 ($P = .04$).

Men were significantly superior ($P = .01$) to women in each of the four conditions of this experiment.

Discussion

The last three experiments reported upon were designed to test the hypothesis that if different uncommon responses are elicited, the tendency to give other uncommon responses will increase and problem solving would be facilitated as a consequence. Partial confirmation of this hypothesis was obtained. Women receiving check lists of uncommon uses tended to perform significantly better than women without this experience in these experiments. Absence of a significant difference between the original and revised experimental groups in Exp. III and IV and the significant superiority of the revised experimental groups over the control groups indicated that the facilitation of problem solving cannot be attributed in any simple fashion to transfer from the check list to the problem-solving situation. The significantly faster solution times of the two experimental groups receiving uncommon uses as compared to the control group that received common uses in Exp. IV suggest in a more direct fashion that it is the uncommonness of the responses elicited that facilitated problem-solving performance.

Although results for the men in the three experiments were in the same direction as the women, they were not statistically significant in most cases. The original experimental group was significantly superior to the control group only in Exp. II and approached significance in Exp. III. The revised experimental groups were not significantly superior to the control groups in either Exp. III or IV. One possible interpretation of the inconsistency in results between men and women is that the problem was too easy for the men. It did not offer sufficient ceiling for them to improve as function of the experimental variables. This is indicated in part by their significantly faster solution times in 8 of the 10 conditions employed in the last three experiments, and a median solution time of approximately 60 sec. for the three experiments.

The initial hypothesis that the elicitation of uncommon responses will facilitate subsequent problem-solving performance involves several assumptions for which there is only fragmentary evidence at the present time. One such assumption is that the tendency to emit uncommon responses is significantly related to facility in solving problems. Experiment I was designed to provide evidence on this point. The results of this study suggest that there is a relationship between the uncommonness of response measure of originality of uses and performance on the two-string problem. This conclusion, however, is only provisional because of the sex differences encountered.

Summary

Four experiments were conducted employing the two-string problem. Experiment I was designed to determine whether originality as measured in terms of uncommonness of response is related to the speed of solution on the two-string problem. Significant but low correlations between these measures were obtained for men but not for women.

Experiments II-IV were designed to determine whether presenting *S*s with lists of uncommon uses for objects would facilitate their subsequent solution of the two-string problem in which these objects may be used. Training of this sort tended to facilitate the problem-solving performance of women in the absence of any obvious specific transfer effects. The direction of the effect was the same among the men, but was not statistically significant in the majority of the conditions.

References

1. Christensen, P. E., Guilford, J. P., & Wilson, R. C. Relation of creative responses to working time and instructions. *J. exp. Psychol.*, 1957, **53**, 82–88.
2. Cofer, C. N. Verbal behavior in relation to reasoning and values. In H. Guetzkow (Ed.), *Groups, leadership and men*. Pittsburgh: Carnegie Press, 1951.
3. Guetzkow, H. An analysis of the operation of set in problem-solving behavior. *J. gen. Psychol.*, 1951, **45**, 219–244.

4. Guilford, J. P. Creativity. *Amer. Psychol.*, 1950, **5**, 444–454.
5. McNemar, O. W. An attempt to differentiate between individuals with high and low reasoning ability. *Amer. J. Psychol.*, 1955, **68**, 20–36.
6. Maier, N. R. F. Reasoning in humans. II. The solution of a problem and its appearance in consciousness. *J. comp. Psychol.*, 1931, **12**, 181–194.
7. Maltzman, I. Thinking: From a behavioristic point of view. *Psychol. Rev.*, 1955, **62**, 275–286.
8. Maltzman, I., Bogartz, W., & Breger, L. A procedure for increasing word association originality and its transfer effects. *J. exp. Psychol.*, 1958, **56**, 392–398.
9. Osgood, C. E. *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.
10. Ross, R. T. Optimum orders for the presentation of pairs in the method of paired comparisons. *J. educ. Psychol.*, 1934, **25**, 375–382.
11. Taylor, D. W., & McNemar, O. W. Problem solving and thinking. *Ann. Rev. Psychol.*, 1955, **6**, 455–482.
12. Thurstone, L. L. Primary mental abilities. *Psychomet. Monogr.*, 1938, No. 4.
13. Wilson, R. C., Guilford, J. P., & Christensen, P. R. The measurement of individual differences in originality. *Psychol. Bull.*, 1953, **50**, 362–370.

(Received January 2, 1958)

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

HUMAN DELAYED-REWARD LEARNING WITH DIFFERENT LENGTHS OF TASK [1,2]

CLYDE E. NOBLE AND WAYNE T. ALCOCK

Montana State University *Louisiana State University*

There is a common belief that reinforcement and feedback in human learning, just as in rodent learning, must be undelayed and specific if optimal proficiency is to be achieved. Lorge and Thorndike (**8**) reported, however, that immediate, quantitative knowledge of results produced no better skill in tossing a ball at an unseen target than did lags in feedback up to 6 sec. That experiment, which is contrary to the familiar generalization found in most textbooks, is rarely cited in discussions of delayed-reward learning. To extend their work, Lorge and Thorndike suggested using "more available and definable responses, such as tapping keys, saying words or numbers, etc." (**8,** p. 193).

Several papers appearing between 1935 and 1951 have been reviewed by Ammons (**1**). Because of methodological variations, the evidence is largely inconclusive, but null results predominate among experiments which do not fill the delay interval with other activity. A study by Saltzman (**16**) attempted to investigate delayed information in verbal learning, but the presence of confounding factors in the procedure makes it difficult to interpret. Unfortunately, the same appears to be true of recent experiments on line drawing (**5, 17**). With only one exception (**2**), researchers have been unable to keep trial spacing, work decrement, or forgetting from complicating their results.

Regardless of one's evaluation of the available data on human *Ss*, however, it is logical to expect that between zero delay (immediate knowledge or reward) and infinite delay (no knowledge or reward) temporal gradients of some kind must exist which are contingent upon discoverable experimental conditions. Recently Bourne (**4**) found increasing errors up to 8 sec. in a multiple-choice concept-identification task where the time interval preceding correct information began after a complex stimulus pattern was removed. Neither a contiguity nor a reinforcement theory would have difficulty explaining the decremental influence of this type of delayed feedback.

Less ambiguous than the human research are animal studies employing delayed reward in instrumental conditioning devices similar to the Skinner box. As interpreted by Spence (**18**), the current evidence indicates that reinforcement effects in rats are undiminished over temporal lags of

[1] This report is based on work done under ARDC Project No. 7707, Task No. 77133, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

[2] The data were collected in 1953 by J. P. Christopher, J. B. Downs, and J. H. Harrell under the supervision of R. B. Card. Statistical analyses were performed by R. L. Frye, Jr., Patricia A. Ray, J. L. Regan, and A. Taylor with the support of the L.S.U. Council on Research. A portion of the paper was read before the Southeastern Psychological Association, March 1957. For helpful comments we are grateful to R. B. Ammons, E. A. and I. McD. Bilodeau, F. M. du Mas, and D. W. Norton.

several seconds, unless receptor and postural orientation to the relevant cues is disrupted. The data show, in fact, that *S*'s ability to maintain a favorable "set" is partly dependent upon the nature of the task. With human *S*s, the importance of task factors shows up in the differential feedback effects in discrete responding and in continuous tracking, as Bilodeau (**3**) has nicely demonstrated. Whether reward or information is withheld seems to be of less consequence than what *S* does during the time interval between response and aftereffect.

Among the task factors which may be relevant to serial multiple-choice learning by delayed reward in man are the number of available responses (N_R) and the number of choice points (N_P) in the problem (**9, 10, 11, 12**); i.e., *task complexity*. By covarying these factors as in a prior experiment (**12**) in this series, one might first determine the relationship between task length (N_P) and time of delay (T_G). If T_G should interact with N_P, then further analysis of the role of N_R holding N_P constant would be indicated. Should N_P and T_G be independent, on the other hand, investigation of N_R would be unnecessary since $N_R = N_P$ for each task. The former type of investigation, therefore, should have priority over the latter. Our basic question is whether increasing the complexity of the task will produce impaired performance over relatively short, unfilled delays of reward. Longer delays, unfortunately, require confounding alterations in the distribution of practice if the reward is to occur contiguously with the relevant stimulus.

The present experiment will study 4-link vs. 10-link compound trial-and-error learning under six delays of serial reinforcement from 0 to 3 sec. Of major interest will be (*a*) the influence of amount of practice (N) on each task, (*b*) the main effects of T_G in the two situations, (*c*) whether T_G interacts with N or N_P, and (*d*) the acquisition curves of the two tasks expressed in terms of response probability (R_p).

Method

Apparatus.—The learning device, described in detail elsewhere (**10, 14**), is a semiautomatic 19-key push-button apparatus called the Selective Mathometer. Pacing stimuli are presented serially by an electronically controlled slide projector, to which *S* responds by selecting keys on a semicircular panel. The stimuli consisted of distinctive electrical circuit symbols (**11**), 4 of which were used in the shorter problem and 10 in the longer one. Stimulus presentation followed an invariant sequence for any given subgroup of four *S*s, but the stimuli appeared in 12 random permutations for each experimental (T_G) group as a whole ($n = 48$).

The *S*'s problem was to discover the correct sequence of reactions corresponding to the 4 or 10 choice points in the two tasks. Like the stimulus series, there were 12 different response permutations for each T_G group, four *S*s having a given permutation. The same keys were exposed as in corresponding conditions of the experiment on the N_P factor (**11**). Considering *S*'s keyboard as being numbered from left to right, the following keys were unmasked in the two problems:

N_P	*Available Keys*
4	4, 8, 12, 16
10	2, 4, 6, 8, 9, 11, 12, 14, 16, 18

Secondary reward (G) is provided by a 3-w green lamp mounted on the reaction panel just below the projection screen. This feedback or reinforcing event, which is independently variable in delay and duration, occurs automatically following each correct reaction ($R+$). Since the noncorrection procedure was employed, trial scores consisted simply of counter records and the polygraph was disconnected.

Procedure.—The experimental plan involved two mixed factorial designs (**7**, Type I), in which each *S*'s problem was to practice an invariant sequence of either 4 or 10 S-R connections by serial reinforcement for 20 trials under one of the following delays of reward: .00, .25, .50, 1.00, 2.00, or 3.00 sec.

In keeping with the distinction between ordinary delayed-reward learning (**18**) and the interpolated-response feedback method (**3, 8**), in which scores are reported one or more trials in arrears, the noncorrection procedure was used as in a recent study of the locus-of-reward factor (**15**). Thus *S* was required to make one and only one response at each choice point, and the locus of reward was serial. Following our conventional procedure (**9, 10, 11, 12, 15**), *S* was instructed to use the forefinger of his preferred hand, to respond as quickly as possible at each choice point, to return to the fixation plate after each choice, and to refrain from responding when the screen was dark. This entails what Bilodeau and Bilodeau (**2**) have called "hand-resting" rather than "hand-maintaining" activity.

To ensure adequate time for signaling a correct response to *S* at the longer delays, the typical 2-sec. stimulus duration was increased to 4 sec. for this experiment. Generally all groups received *G* for an *R*+ before the dark period began, but of necessity the time between reward onset and stimulus offset tended to vary inversely with T_G. Reward duration was .75 sec. and the interstimulus period was 1.44 sec. as usual, but the intertrial interval lasted for 10.88 sec. rather than 6.88 sec. This resulted in more distributed practice than formerly (e.g., **11**), but had the advantage of maintaining a constant distribution of stimuli and trials for all *S*s under the present conditions (cf. **10**, p. 391).

The *S*s were 576 male and female recruits at Lackland Air Force Base. They were divided without bias into six groups of 48 *S*s within each of the two tasks, and the 12 S-R permutations were randomly assigned to each T_G group independently. The two problems were run concurrently. All *S*s were tested in subgroups of four on a multiple-unit version of the Selective Mathometer described earlier (**11**). Following the experiment, each *S* was isolated from the remaining *S*s to prevent collusion.

TABLE 1

ANALYSIS OF VARIANCE OF CORRECT RESPONSES (*R*+) FOR SIX DELAY-OF-REWARD GROUPS IN 4-LINK PROBLEM DURING TRIALS 1–10

Source	*df*	*MS*	*F*
Between *S*s	287		
Delay of reward (T_G)	5	6.15	1.04*
Error (*b*)	282	5.94	
Within *S*s	2592		
Trials (*N*)	9	209.16	282.30**
Interaction ($N \times T_G$)	45	.48	.65
Error (*w*)	2538	.74	
Total	2879		

* $P > .20$.
** $P < .001$.

TABLE 2

MEAN PERCENTAGE CORRECT RESPONSES (*R*%) DURING TRIALS 1–10

Number of Choice Points (N_P)	Delay of Reward (T_G) in Seconds					
	0	1/4	1/2	1	2	3
4	69.90	69.74	71.30	72.86	71.35	64.69
10	25.67	25.60	27.33	28.50	23.77	22.77

RESULTS

The 4-link problem.—To evaluate the amount-of-practice (*N*) and delay-of-reward (T_G) factors in the 4/4 situation, a 10×6 analysis of variance was performed upon the *R*+ scores of the six delay groups during Trials 1–10 ($n = 48$). Due to the fact that *S*s reached a performance level over 90% of mastery by Trial 10, the possibility exists that heterogeneity of form and variance might invalidate the *F* test beyond that point (**7, 12, 15**). The analysis, therefore, was carried out upon the first half of practice only. As shown in Table 1, the main effect of *N* is significant, but neither the T_G factor nor the $N \times T_G$ interaction reaches the 5% level. The upper half of Table 2 records the mean percentage of correct choices obtained during this period by each group. There is no evidence of any trend due to T_G.

The 10-link problem.—To evaluate the *N* and T_G factors in the 10/10 situation, the same type of analysis of variance was performed as above. In order to distribute the performance scores over a similar range and to use the same degrees of freedom, the

TABLE 3

ANALYSIS OF VARIANCE OF CORRECT RESPONSES ($R+$) FOR SIX DELAY-OF-REWARD GROUPS IN 10-LINK PROBLEM DURING TRIALS 11–20

Source	*df*	*MS*	*F*
Between *Ss*	287		
Delay of reward (T_G)	5	82.35	1.62*
Error (b)	282	50.97	
Within *Ss*	2592		
Trials (N)	9	245.58	189.52***
Interaction ($N \times T_G$)	45	1.80	1.39**
Error (w)	2538	1.30	
Total	2879		

* $.10 < P < .20$.
** $P < .05$.
*** $P < .001$.

analysis was restricted to the second half of practice in this case. Referring to the summary presented in Table 3, the main effect of N is again significant and there is a significant $N \times T_G$ interaction, but the T_G factor itself fails to produce an overall decrement at the 5% level. The probability that the obtained differences are due to chance is about .15.

A significant Practice × Delay interaction means that the various T_G groups do not differ to the same extent on every trial, suggesting the possibility of systematic though minor decreases in rate of change with increasing delay in the 10-link problem. Inspection of the lower half of Table 2, however, does not support this hypothesis. The rank-difference correlation between the T_G values and the $R\%$ scores is only .49, which for $df = 4$ is not significantly greater than zero. It may be noted that the rank order of T_G groups in the 10/10 situation is exactly the same during Trials 11–20 as during Trials 1–10.

Acquisition curves.—In view of the absence of significant differences in the 4-link problem, the trial scores of the six delay groups were pooled to give a general curve based on 288 *Ss*. The same was done for the 10-link problem, despite the observed interaction, then the acquisition curves for the two problems (N_P) were plotted as the probability of a correct response (R_p) over successive trials (N).

In agreement with experiments on response availability (**9, 12**), task length (**11**), and locus of reward (**15**), all curves were asymmetrical and sigmoidal, conforming to the rational function proposed by Noble (**10**):

$$R_p = a(i)^{r^N} \qquad [1]$$

where a = asymptote of 1.00, i = reciprocal of number of choice points ($1/N_P$), r = rate parameter dependent on N_P, and N = number of trials by the noncorrection procedure.

By minimizing the sums of squares of the residuals to determine the parameter r, the best-fitting equations are as follows:

$$N_P = 4, \quad T_G = 0\text{–}3 \text{ sec.}: \quad R_p = 1.00\ (.250)^{.747^N} \qquad [2]$$

$$N_P = 10, \quad T_G = 0\text{–}3 \text{ sec.}: \quad R_p = 1.00\ (.100)^{.918^N} \qquad [3]$$

The percentages of variance accounted for are 99.3% and 99.8% for Equations 2 and 3, the inflection points occurring at 1.12 and 9.50 trials, respectively. These computational procedures have been fully described elsewhere (**11**). It is important to observe that both curves rise faster (have smaller r parameters) than their counterparts in the earlier experiment (**11**) on N_P. This difference is probably due to the use of more distributed practice and the non-

correction procedure in the present study.[3]

Discussion

Delaying serial reward up to 3 sec. has no significant effect upon human trial-and-error learning, regardless of length of task or amount of practice. Our negative findings thus agree with those of Lorge and Thorndike (**8**). As they pointed out, ". . . it is almost impossible to prevent human subjects from keeping a connection . . . in mind until the after-effect occurs or from recalling it to mind when the after-effect occurs and then repeating some equivalent of the after-effect. The real relation would then be immediate, regardless of how long an ostensible delay we used" (**8**, pp. 193–194).

From the quotation above, it is evident that the Thorndikean position on delayed-reward learning is very close to that of other S-R behavior theorists such as Hull (**6**) and Spence (**18**). Without some kind of competing interpolated responses, the occurrence of "sheer empty delay" (**8**, p. 186) does not necessarily lead to impaired performance. The hypothetical process generally thought to mediate this time-bridging ability of the organism is secondary reinforcement. Thus, Hull (**6**) suggested that temporal lags in reward following a critical response can maintain habit strength (H) through exteroceptive and proprioceptive stimuli having acquired reward value from association with primary goal events. Originally Hull had proposed a brief negatively accelerated delay function of the order of 30 sec. (gradient of reinforcement) operating in conjunction with secondary reinforcement to generate more permanent effects (goal gradient). As investigators gained further control over reward factors, Spence (**18**, p. 150) later argued that all empirical gradients may result from the stimulus generalization of immediate secondary reinforcement. More recently, however, Spence (**18**, p. 151) has forsaken habit-strength interpretations in favor of simple experimental extinction, thereby adopting a Guthrian type of counterconditioning hypothesis.

But these assumptions were developed to account for the behavior of animals working for primary rewards. It remains to be seen whether, in human learning by secondary reinforcement, decremental effects of delay of reward can be subsumed under the principle of experimental extinction. Until studies of multiple-choice learning are performed which employ a greater range of T_G values with and without disrupting interpolated activity, this question cannot be answered. The writers doubt that any published human data contain unequivocal evidence of impairment within this range due to "sheer empty delay" (**8**). A similar point has been made cogently with reference to the acquisition of discrete positioning acts in a recent paper by Bilodeau and Bilodeau (**2**). Five experiments (varying delay of feedback, time between feedback and the next response, and the intertrial interval up to 7 days) failed to show any effects other than those due to distribution of practice. The only relevant data not in agreement with this conclusion are those of Bourne (**4**) but, as he pointed out, a (perceptual) forgetting hypothesis is as tenable as a (motor) reinforcement-generalization hypothesis.

Summary

This experiment compared 4-link and 10-link human trial-and-error learning under six delays of serial reinforcement from 0 to 3 sec. The *Ss* ($n = 48$ for each of 12 conditions) practiced on a random sequence of either 4 or 10 choice points (N_P) for 20 trials by the noncorrection procedure. Delay intervals were free of interpolated activity.

The shorter task produced very rapid acquisition, but there were no significant differences due either to delay of reward (T_G) or to interaction with amount of practice (N). The longer task resulted in much slower acquisition and there was a significant $N \times T_G$ interaction late in practice, but the effect of T_G itself only

[3] Theoretical and experimental support for the latter half of this statement has been presented elsewhere (**13**).

reached the 15% level. Response probability curves confirmed previous data on the influence of task length and its interaction with amount of practice.

It was concluded that temporal lags in serial reward up to 3 sec. have no significant effect per se upon human trial-and-error learning, regardless of task complexity. Implications of the results for current explanations of delayed-reward and delayed-feedback phenomena were discussed.

References

1. Ammons, R. B. Effects of knowledge of performance: A survey and tentative theoretical formulation. *J. gen. Psychol.*, 1956, **54**, 279–299.
2. Bilodeau, E. A., & Bilodeau, I. McD. Variation of temporal intervals among critical events in five studies of knowledge of results. *J. exp. Psychol.*, 1958, **55**, 603–612.
3. Bilodeau, I. McD. Accuracy of a simple positioning response with variation in the number of trials by which knowledge of results is delayed. *Amer. J. Psychol.*, 1956, **69**, 434–437.
4. Bourne, L. E., Jr. Effects of delay of information feedback and task complexity on the identification of concepts. *J. exp. Psychol.*, 1957, **54**, 201–207.
5. Greenspoon, J., & Foreman, S. Effect of delay of knowledge of results on learning a motor task. *J. exp. Psychol.*, 1956, **51**, 226–228.
6. Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
7. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.
8. Lorge, I., & Thorndike, E. L. The influence of delay in the after-effect of a connection. *J. exp. Psychol.*, 1935, **18**, 186–194.
9. Noble, C. E. Compound trial-and-error learning as a function of response availability (N_R). *J. exp. Psychol.*, 1955, **49**, 93–96.
10. Noble, C. E. Human trial-and-error learning. *Psychol. Rep.*, 1957, **3**, 377–398.
11. Noble, C. E. The length-difficulty relationship in compound trial-and-error learning. *J. exp. Psychol.*, 1957, **54**, 246–252.
12. Noble, C. E. Supplementary report: Response availability in human trial-and-error learning. *J. exp. Psychol.*, 1957, **53**, 359–360.
13. Noble, C. E., Alcock, W. T., & Noble, J. L. The interaction of type of choice procedure with amount of practice in trial-and-error learning under two reward conditions. *J. Psychol.*, 1958, **46**, 295–301.
14. Noble, C. E., & Farese, F. J. An apparatus for research in human selective learning. *J. Psychol.*, 1955, **39**, 475–484.
15. Noble, C. E., & Noble, J. L. Human trial-and-error learning under joint variation of locus of reward and type of pacing. *J. exp. Psychol.*, 1958, **56**, 103–109.
16. Saltzman, I. J. Delay of reward and human verbal learning. *J. exp. Psychol.*, 1951, **41**, 437–439.
17. Saltzman, I. J., Kanfer, F. H., & Greenspoon, J. Delay of reward and human motor learning. *Psychol. Rep.*, 1955, **1**, 139–142.
18. Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

(Received January 2, 1958)

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

CONCEPT IDENTIFICATION AS A FUNCTION OF COMPLETENESS AND PROBABILITY OF INFORMATION FEEDBACK [1,2]

LYLE E. BOURNE, JR., AND R. BRIAN PENDLETON

University of Utah

This paper reports two studies aimed at examining performance in a concept identification situation when the learner is provided with (*a*) complete vs. incomplete feedback, and (*b*) several degrees of intermittent information feedback. These experiments represent two in a program, the ultimate aim of which is to elucidate those circumstances wherein a learner makes maximal use of the information provided in a problematic situation.

The first investigation in this series (**2**) demonstrated a strict dependence of concept identification performance on delay of information feedback. The number of errors made by *S* in solving this type of problem increased at a positively accelerated rate with delay. It is plausible to expect performance to depend in a similar fashion on completeness and probability of feedback. There is little evidence upon which to base a prediction concerning the effects of completeness of feedback except by comparison of two independent studies wherein *S* was allowed either to respond to a stimulus until correct (**1**) or to respond once, being subsequently informed as to the correctness of the response (**2**). On the other hand, there have been recently many attempts to determine the effects of partial reinforcement on performance (**4**, **5**, **6**, **7**, **8**, **9**, **10**). Clearly, there are differences between the reinforcement used in these studies and the type of information feedback employed in the present experiments. However, the results suggest that intermittent feedback maintains a strong influence over behavior in a multiple-choice situation.

Restle (**11**, **12**) has proposed a mathematical theory of discrimination learning in which the stimulus situation is described as a set of elements or cues that might be either relevant or irrelevant to a particular response. By assuming (*a*) that relevant cues are conditioned to a response, (*b*) that irrelevant cues are adapted or rendered nonfunctional during learning, and (*c*) that the probability of a response is proportional to the cues adapted to it, some success in predicting the results of two-choice discrimination learning problems was obtained. In addition, by assuming that the degree to which a cue is differentially reinforced determines its "validity," the model offered a descriptive account of how learning rate may depend upon reinforcement probability. Restle's equation (**12**, Equation 8), for validity as a function of the proportion of reinforcements, has yielded strikingly accurate predictions for a two-choice situation where one or the other response is reinforced on every trial.

If reinforcement probability may be interpreted as the proportion of trials on which information feedback is presented, certain quantitative predictions concerning the role of this variable in a four-choice problem can be made. Walker (**13**) has generalized and modified Restle's framework to make it applicable to a concept identification situation. He assumed that the number of relevant

[1] This research was supported in part by a grant from the Research Committee, University of Utah, and in part by a research grant (M-1759) from the National Institute of Mental Health, U. S. Public Health Service.

[2] The authors wish to thank Clinton M. Walker and David L. Brown for their help in collecting the data reported.

and irrelevant cues in a discrimination problem is a function of the amount of relevant and irrelevant information in the stimulus patterns. Then, using Restle's equation (**11**, Equation 8) for approximate number of errors, he found a close relationship between predicted and obtained concept identification performance. This type of analysis, with a further modification of the validity formula so that cue validity is 1 for 100% feedback and 0 for 0% feedback, will be used to compare predicted and obtained errors in the data to be reported.

General Method

Task.—As in the earlier concept identification studies (**1, 2, 3**), *S* was presented with a series of geometric patterns each of which was a combination of the levels of x relevant and y irrelevant stimulus dimensions. To each pattern, *S* responds by pressing one of four keys to identify the category to which the pattern belongs. The four keys correspond to the four possible combinations of the two two-level dimensions which were relevant to problem solution. A dimension is relevant to solution if it is necessarily used in the correct identification of the patterns. If a dimension is irrelevant, it appears at each of its two levels within the stimulus patterns but cannot be consistently used to correctly classify the patterns. For a more complete discussion of the task, see Bourne (**2**).

The criterion of problem solution was 16 consecutively correct identifications. An analysis of data from earlier studies for which the criterion was 32 correct responses indicated that less than 1% of *S*s reaching 16 correct failed to continue without error to 32. All tasks were self-paced in that *S* was allowed as much time as needed to respond to any pattern.

Apparatus.—The stimulus patterns were presented visually to *S* with a Dunning Animatic filmstrip projector. The *S* sat at a distance of 2 ft. from the projection screen with the response keyboard before him. The keys on *S*'s board activated lights on a control panel located behind a partition and out of *S*'s view. Thus, *E* could record each response and in turn provide feedback to *S* as to the correctness of each response. For all problems, the over-all time interval between *S*'s response and the appearance of the next pattern was 5 sec. During that interval, the screen upon which the pattern had been presented was blank. Immediately upon *S*'s response to any pattern, *E* provided information feedback (if it was scheduled) by means of a panel of lights mounted on the response unit. The duration of feedback was always 2 sec.

Stimulus lists.—Nine filmstrip series were prepared to present problems at three degrees of task complexity. Three different problems, defined in terms of the two dimensions relevant to solution, were used: (*a*) orientation (upright or tilted figures) and form (square or triangle), (*b*) vertical position (figures at the top or bottom of screen) and size (large or small figures), and (*c*) color (red or green figures) and number (one or two figures on the screen in the same pattern). Task complexity was graduated in terms of the number of dimensions employed which were irrelevant to solution. The degrees of complexity selected were 1, 3, and 5 irrelevant dimensions. When a dimension was neither relevant nor irrelevant, it appeared at only one of its levels within the patterns.

Since each pattern could take on one of the two levels of any relevant or irrelevant dimension, the amount of information in the stimuli, assuming the two levels of a dimension to be completely discriminable, could be quantified in bits of information by evaluating $\log_2$ of the number of equally probable alternative stimuli in the problem being presented. Thus, in these problems, there were 2 bits of relevant information and either 1, 3, or 5 bits of irrelevant information.

Experiment I

The purpose of the first study was to investigate the effect of variation in feedback completeness upon the acquisition of concepts.

Procedure

Subjects.—Fifty-four students in elementary psychology courses were each assigned randomly to solve one of the problems in one of six treatment combinations. Each *S* was presented at the outset with detailed instructions as to the nature of his task, the operation of the controls, the meaning of the information lights, and the criterion of problem solution.

Design.—A $3 \times 3 \times 2$ factorial design was used with three levels of task complexity (irrelevant information), three different problems (dimensions relevant to solution), and two degrees of feedback completeness. Feedback

was effected by one of two interchangeable light panels placed directly above and about 4 in. from the four keys on the response unit. One panel, consisting of two jeweled lamps, provided simple right-wrong feedback subsequent to *S*'s response. This procedure was identified as incomplete. The second panel, consisting of four jeweled lamps, one directly above each of the four response keys, provided a greater degree of feedback; the light above the correct key for any given pattern was lighted subsequent to *S*'s response to that pattern. This procedure was identified as complete.

Results

Two response measures were used: trials to solution and number of errors. Due to the high correlation between these measures ($r = .901$) and since the instructions to *S* stressed accuracy in problem solving, number of errors was considered the main dependent variable. Figure 1 presents graphically the mean number of errors made by the complete feedback and the incomplete feedback groups as a function of amount of irrelevant information within the patterns. An analysis of variance on these data is summarized in Table 1. A Bartlett's test for homogeneity of variance on the residual error term indicated no significant heterogeneity ($\chi^2 = 18.91$, 17 *df*, $P > .30$). The

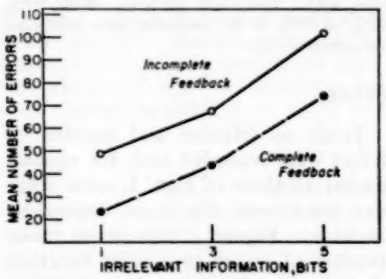

FIG. 1. Mean number of errors for the complete feedback and the incomplete feedback groups as a function of the amount of irrelevant information. Each plotted point represents the data from nine *S*s.

TABLE 1

ANALYSIS OF VARIANCE OF NUMBER OF ERRORS, EXP. I

Source of Variance	*df*	*F*
Completeness of feedback (C)	1	4.33*
Irrelevant information (I)	2	11.46**
Linear	1	22.47**
Quadratic	1	.45
Problems (P)	2	3.98*
C × I	2	.15
C × P	2	1.09
I × P	4	.90
C × I × P	4	.48
		Residual *MS*
Residual	34	1487.98
Total	53	

* $P < .05$.
** $P < .01$.

main effects of completeness of information feedback, amount of irrelevant information, and problems were statistically significant. The significance of the irrelevant information source of variance is consistent with all earlier results (**1, 2, 3**) and indicates that the number of errors made in solving this type of problem is an increasing function of task complexity. Figure 1 indicates that this increase is a nearly straight line function of information in bits. In an orthogonal polynomial analysis applied to the function, only the linear term reached the .01 level of significance. The near constancy of the difference between the complete feedback and the incomplete feedback *S*s over the range of task complexity used, as shown in Fig. 1, was supported by the nonsignificance of the completeness of feedback by irrelevant information interaction term of the analysis.

Consistent with earlier results (**2**), the variance identified with problems was statistically significant. The greatest number of errors was associated with Problem *b* which con-

tained a relevant "positional" dimension, i.e., a dimension referring to the position of the figure on the screen rather than to the character of the figure itself.

With trials-to-solution as the per-performance measure, essentially identical results were found; the analysis of variance showed the same sources of variance to be significant.

Walker's modification (13) of the Restle model (11) was made with the assumption of complete feedback to *S* and is not necessarily adequate to account for the data of the incomplete feedback group. The equation applicable for predicting approximate number of errors made is of the form

$$E \simeq k\left\{\tfrac{1}{2}+\tfrac{1}{2}\left[\frac{\log \theta}{(1-\theta)\log(1-\theta)}\right]\right\} \quad [1]$$

where θ is the ratio of relevant to total stimulus information in the problem and k is a proportionality constant. For the three levels of complexity used in the experiment, θ is 2/3, 2/5, and 2/7, respectively. The mean number of errors made by *S*s in the complete feedback group for the three levels of task complexity were 23.00, 42.89, and 73.78. The predicted errors, from Equation 1, for these three conditions are $1.05k$, $1.98k$, and $3.11k$. For convenience of comparison the predictions are reduced to the following scalar-vector product: $1.05k$ (1 1.89 2.96). The obtained errors represented in the same form are: 23.00 (1 1.86 3.21). Solving for k as the ratio of the scalar multipliers, one gets 21.82. The correspondence between the predicted and the obtained vectors, though not perfect, is promising. As was expected, the correspondence between this prediction and the errors observed for the incomplete feedback group was less exact.

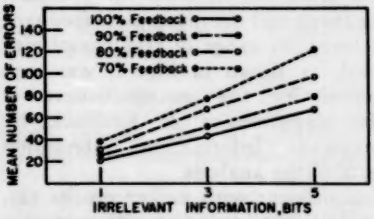

FIG. 2. Mean number of errors as a joint function of the amount of irrelevant information and the probability of feedback. Each plotted point represents the data from nine *S*s.

EXPERIMENT II

Experiment II was designed to assess concept identification performance as determined by various probabilities of complete information feedback.

Procedure

Subjects.—The *S*s were 108 students, assigned randomly and individually to one of 12 main treatment combinations. None had served in Exp. I. The general instructions were essentially the same as those used in Exp. I. In addition, all *S*s were told that, on some trials, feedback might not be presented. The *S*s were to interpret this event as meaningless insofar as the correctness of the response was concerned and to proceed on to the next pattern.

Design.—A 3 × 3 × 4 factorial design was employed with three levels of task complexity, three different problems, and four schedules of feedback probability, 100%, 90%, 80%, and 70%. The schedule of feedback for any given probability level was the same for all *S*s serving in that condition and was determined a priori by reference to a table of random numbers so that, within every 128 patterns, 100%, 90%, 80%, or 70% of *S*'s responses were reinforced (or corrected) by *E*.

Results

Trials to solution and number of errors were recorded and, for reasons similar to those of Exp. I, error score was considered the main dependent variable. Figure 2 represents mean number of errors as a joint function of task complexity and probability of feedback. The analysis of variance performed on these data is summarized in Table 2. A Bartlett's test indicated no significant hetero-

geneity of variance in the residual term of the analysis ($\chi^2 = 34.78$, 35 *df*, $P > .20$). The main effects of amount of irrelevant information and of feedback probability were significant ($P = .01$). The significance of the irrelevant information source is consistent with the results of the first study, and again only the linear term of an orthogonal polynomial analysis of this function reached a significance level greater than .01. In addition, the mean number of errors increased as probability of feedback decreased. An orthogonal polynomial analysis applied to this function suggests that it, too, is linear.

Inconsistent with the results of Exp. I is the failure of the Problems term to reach significance. However, as in the earlier results, Problem *b*, the problem with a relevant "positional" dimension, was the most difficult.

A theoretical analysis of the errors made, following Walker (13), necessitated the introduction of the notion of cue validity. Recall θ, the ratio of relevant to total information. Restle (12) suggested that θ might act in a multiplicative fashion with cue validity, V, so that the equation predicting the number of errors, Equation 1, would involve a term ϕ, in place of θ, which is equal to θV. Restle presented an expression for V in a two-choice situation wherein $V = 1$ when reinforcement of a response is either 0 or 100% and $V = 0$ when reinforcement is 50% for either response. For the four-choice concept identification situation, the function was modified such that $V = 1$ for 100% feedback and 0 for 0%. The equation used in analyzing the present data was of the form

$$V = .5\pi^2 + .5\pi \qquad [2]$$

where π is the proportion of trials accompanied by feedback.

Table 3 shows the result of a theoretical analysis of the error data, including cue validity, theoretical and observed error vectors, and the proportionality constant. The degree of correspondence between observed and expected vectors for the various experimental conditions was extremely close, the greatest discrepancies appearing in

TABLE 2

ANALYSIS OF VARIANCE OF NUMBER OF ERRORS, EXP. II

Source of Variance	*df*	*F*
Feedback probability (F)	3	5.87*
Linear	1	17.47*
Quadratic	1	.14
Cubic	1	.01
Irrelevant information (I)	2	31.64*
Linear	1	63.27*
Quadratic	1	.01
Problems (P)	2	.66
F × I	6	.58
F × P	6	.78
I × P	4	.22
F × I × P	12	.66
		Residual *MS*
Residual	72	1073.06
Total	107	

* $P < .01$.

TABLE 3

THEORETICAL ANALYSIS OF ERRORS

Group	*V*	Observed Vector[a]	Theoretical Vector	*k*
100%	1	22.56 (1 1.89 2.96)	1.054*k* (1 1.89 2.96)	21.40
90%	.855	26.78 (1 1.94 2.98)	1.272*k* (1 1.92 3.00)	21.05
80%	.720	31.00 (1 2.19 3.15)	1.578*k* (1 1.95 3.07)	19.65
70%	.595	39.23 (1 1.96 3.08)	2.008*k* (1 1.97 3.11)	19.54

[a] To obtain the actual mean number of errors made by *Ss* in the 1-, 3-, and 5-bit irrelevant conditions at each level of feedback probability, multiply the scalar by each of the three entries in the vector.

TABLE 4

DIFFERENCE IN MEAN NUMBER OF TRIALS ON WHICH FEEDBACK WAS PRESENTED BETWEEN THE 100% CONDITION AND ALL OTHERS

(16 criterion trials excluded)

Task Complexity	Feedback Probability			Mean Diff.
	90%–100%	80%–100%	70%–100%	
1 bit	−1.4	−3.8	−1.9	−2.4
3 bits	3.3	−3.8	20.6	6.7
5 bits	6.3	23.2	31.4	20.3
Mean	2.7	5.2	16.7	

the 80% feedback condition. Note also that k was roughly the same for all conditions and compared well with k estimated from the data of Exp. I. There was, however, some tendency for a systematic decrease in k with the reduction in feedback.

A final analysis made on the data was suggested by the possibility that the actual number of *reinforced* (feedback) trials might be the same for all feedback conditions. If *S* simply disregarded those trials on which no feedback was presented, the increase in number of trials required as feedback was reduced might be a reflection of the necessity for a critical number of reinforced trials to reach the solution of any problem. Table 4 presents the difference in mean number of trials on which information was presented between the 100% group and all others. Note the increasingly inhibiting effect of reduced feedback as problem complexity was increased. This finding suggests an interaction of these two variables although there was no statistical support for this conclusion. Actually fewer trials accompanied by feedback were required by *S*s in the 1-bit irrelevant groups whereas more trials were needed by the 3- and 5-bit irrelevant groups.

DISCUSSION

The results of these studies clearly support several of the proposed hypotheses. As in most earlier studies (**1, 2, 3**), task difficulty was found to be effectively controlled by varying the amount of irrelevant information in the patterns. Performance was an inverse linear function of irrelevant information. Both completeness of information feedback and probability of information feedback were consistently significant sources of variance in the analyses. The data support the notion that reducing completeness of feedback simply reduces performance by a constant amount, regardless of task complexity. This factor, then, cannot account for the difference in *shape* of the performance-complexity function evident from a comparison of the results of two previous experiments (**1, 2**). It is plausible, then, to attribute the discrepancy in these earlier data to another important methodological difference between them—the sophistication of *S*s. In the first of these experiments, *S*s served in several sessions on consecutive days, whereas in the other, all *S*s were naive and served only once.

The analyses of the effect of probability of feedback suggest that its effect depends upon task complexity, reduced probability being more inhibiting as task complexity increases. There is, however, no statistical support for this conclusion. It is plausible to argue that, had several other schedules of feedback or several other levels of irrelevant information been introduced into the experimental design, a significant interaction might have been found. However, increasing these effects would lead to a severe sampling problem since the data of pilot studies indicate that sizable percentages of *S*s fail problems of greater complexity or less feedback probability.

Of obvious significance is the success obtained with the theoretical analysis of the present data. Coupling the general framework of Restle (**11, 12**) with the modifications suggested by Walker (**13**), quantitative predictions concerning the way in which performance (errors) in a four-choice problem should be affected by changes in irrelevant stimulus information and intermittent

information feedback were made and, by and large, confirmed. The close fits of the predictions to the empirical results clearly support the assumption that a functional relationship exists between the ratio of relevant to total stimulus cues, as described by Restle, and the ratio of relevant to total information in bits, within the stimulus patterns, as discussed by Walker. The extention of the concept of cue validity, beyond its application to reinforcement probability in two-choice discrimination learning, was also confirmed. The data indicated that validity was affected by the proportion of trials in a four-choice problem accompanied by complete information feedback; in general, an inverse relationship exists between feedback probability and cue validity. According to Walker, k should remain constant regardless of feedback or complexity conditions. Although there is no great variability in the estimated k's, there appears to be a systematic reduction in the k value as feedback probability is reduced. The reason for this trend is not obvious but may be due to the chosen form of the validity equation. Other constants were tried in Equation 2 but these failed to change the trend in k. Because of the close fits obtained, further information should be available before any modification is made in the form of Equation 2. It should be stressed that the analysis performed is not represented as strictly in accord with Restle's probability model. The data, however, were encouraging enough to warrant a continuation of the conceptions used.

Summary

Two experiments were conducted to test the effects of completeness and probability of information feedback upon concept identification. In Exp. I, 54 *Ss* served in a factorial design with three levels of task complexity (1, 3, and 5 bits of irrelevant information within the stimulus patterns) and two degrees of feedback completeness (a complete feedback and an incomplete feedback condition). The task for *S* was to learn the correct method of classifying visually presented stimulus patterns.

In Exp. II, 108 *Ss* served in a factorial design with the levels of task complexity used in Exp. I and with four different schedules of feedback probability, 100%, 90%, 80%, and 70%. The task was essentially the same in the two studies.

The major findings were: (*a*) Feedback completeness significantly affected concept identification. There was a relatively constant difference between the incomplete and the complete feedback groups in performance at all levels of task complexity. (*b*) Feedback probability also was a consistent source of variance in the analyses. There was an inverse linear relationship between errors and feedback probability. Although the analysis of variance indicated no significant Probability × Complexity interaction, the data do suggest this as a possibility. (*c*) Varying the amount of irrelevant information in the stimulus patterns was an effective method of manipulating task difficulty. In both studies, performance decreased linearly with increased irrelevant information. (*d*) The performance observed under various conditions of feedback probability was adequately predicted by Walker's (**13**) modification of Restle's (**11, 12**) model for discrimination learning.

References

1. Archer, E. J., Bourne, L. E., Jr., & Brown, F. G. Concept identification as a function of irrelevant information and instructions. *J. exp. Psychol.*, 1955, **49**, 153–164.
2. Bourne, L. E., Jr. Effects of delay of information feedback and task complexity on the identification of concepts. *J. exp. Psychol.*, 1957, **54**, 201–207.
3. Brown, F. G., & Archer, E. J. Concept identification as a function of task complexity and distribution of practice. *J. exp. Psychol.*, 1956, **52**, 316–321.
4. Goodnow, J. G., & Postman, I. Probability learning in a problem-solving situation. *J. exp. Psychol.*, 1955, **49**, 16–22.
5. Grant, D. A., Hake, H. W., & Hornseth, J. P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, **42**, 1–5.
6. Green, E. J. Concept formation: A problem in human operant conditioning. *J. exp. Psychol.*, 1955, **49**, 175–180.
7. Humphreys, L. G. Acquisition and extinction of verbal expectancies in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, **25**, 294–301.

8. HUMPHREYS, L. G. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J. exp. Psychol.*, 1939, **25**, 141–158.
9. KANFER, F. H. The effect of partial reinforcement on the acquisition and extinction of a class of verbal responses. *J. exp. Psychol.*, 1954, **48**, 424–432.
10. McNAIR, D. M. Reinforcement of verbal behavior. *J. exp. Psychol.*, 1957, **53**, 40–46.
11. RESTLE, F. A theory of discrimination learning. *Psychol. Rev.*, 1955, **62**, 11–19.
12. RESTLE, F. Theory of selective learning with probable reinforcements. *Psychol. Rev.*, 1957, **64**, 182–191.
13. WALKER, C. M. Concept identification as a function of amounts of relevant and irrelevant information. Unpublished doctor's dissertation, Univer. of Utah, 1958.

(Received January 2, 1958)

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

TRANSFER EFFECTS ON A ROTARY PURSUIT TASK AS A FUNCTION OF FIRST-TASK DIFFICULTY[1]

DANIEL S. LORDAHL AND E. JAMES ARCHER

University of Wisconsin

In his review of the problem of task difficulty and transfer of training, Day (**3**) pointed out that many methods have been used to vary the difficulty of the tasks employed, and that it has been hard to generalize about the results. Day concluded that one of the principal problems in this area is the specification of the mode of varying task difficulty. More detailed knowledge of how performance is affected when *S*s transfer from one level of difficulty to another may aid in the specification of relevant dimensions of difficulty.

Ammons, Ammons, and Morgan (**1**) studied the effects of transfer between tasks which were of unequal difficulty. Using a rotary pursuit task, they examined the effects of four speeds of rotation (40, 50, 60, and 70 rpm) before and after a rest period. They also varied the degree of distribution of practice in a factorial design. Under all conditions, the duration of initial practice was 5 min. The basic results were: (*a*) at all stages of practice, greater time on target was associated with slower target speeds, (*b*) there was no interaction between distribution of practice and target speed, and (*c*) shifting from slower to faster speeds resulted in better performance than shifting from faster to slower, although practice at the same speed on both tasks was better than other conditions in most cases. One limitation of this study, as the authors pointed out, was that the degree of first task practice was relatively low.

In the present study, longer periods of practice on the first and second tasks were used than in the Ammons, Ammons, and Morgan study. It was thought that such a practice schedule would increase the transfer effects and permit a more extensive study of these effects.

In addition, two dimensions of task difficulty were studied: (*a*) speed of target rotation, which has been shown to produce different levels of performance (**1, 6**); and (*b*) radius of the target orbit, which was expected to affect learning and transfer since different target orbit radii require different amounts of physical movement. Since there was no interest in the interaction of the two methods of varying difficulty, a separate experiment was performed for each variable, holding the other constant.

PROCEDURE

Subjects.—Three groups of 15 women students in introductory psychology courses at the University of Wisconsin served individually in each of the two experiments (total $N = 90$). They volunteered for the experiment, were given extra credit towards their course work for serving, and no *S* was permitted to serve in both experiments. Each *S* was naive to the task and apparatus and was assigned randomly to one of the six subgroups. Under all conditions, *S* was required to serve for two sessions, on two consecutive days at about the same hour. Four additional *S*s were run but their data were discarded. Three of these did not return for the second session and the fourth was dropped because of a power failure during the experiment.

Apparatus and task.—A variable speed-

[1] This research was supported by a grant from the Research Committee from funds provided by the Wisconsin Alumni Research Foundation. The writers are also indebted to the Numerical Analysis Laboratory for its assistance in the statistical computations.

variable radius pursuit rotor was used. The length of the stylus from the hinge to the point was 6 in. and the point of the stylus contained a ball bearing to reduce friction and maintain a constant-sized stylus tip.

By means of interchangeable turntables *E* varied the distance of the target from the center of the turntable. These target orbit radii were 2.0, 3.5, and 5.0 in. (center to center). The target was a brass disc of .75 in. diameter and was set flush with the surface of the turntable.

The turntable could be adjusted to run at speeds of approximately 15–85 rpm. Using a stroboscopic attachment, *E* adjusted the speed to within .5 rpm of the desired values of 40, 60, and 80 rpm.

When the stylus point contacted the target, the circuit was completed to the Tracking-time Analyzer. This apparatus has been described elsewhere (2), its essential features being a bank of 13 resettable counters and a .01-sec. scale Standard Electric clock. The clock was used to obtain total time-on-target measures. The first of the counters directly recorded each contact between the stylus point and the target. The other counters were activated only after *S* had been continuously on target for various minimum durations. The durations measured by these counters were .00, .05, .10, .20, .30, .40, .60, .80, 1.00, 1.20, 1.60, 2.00, and 5.00 sec.

The *S* could score only during the trial, although the rotor was allowed to run continuously throughout each session.

The sounds from the operation of the relays and counters were clearly audible to *S* and allowed for limited knowledge of results. However, the Tracking-time Analyzer and *E* were separated from *S* by a partial wall and *S*'s scores were not made known to her.

Conditions.—All *S*s received thirty 30-sec. trials per day, with a 60-sec. rest interpolated between trials. During the intertrial rest, *S* placed the point of the stylus in the center of the turntable and read magazines on a nearby desk.

The "speed experiment" studied the effect of transferring from one of three different speeds of target rotation and the "radius experiment" studied the effect of transferring from one of three different target orbit radii.

Each of the three groups in the speed experiment practiced tracking the target at one of three speeds on Day 1: 40, 60, or 80 rpm. On Day 2 all *S*s in this experiment practiced tracking a 60-rpm target. On both days the target orbit radius was held constant at 5.0 in. These groups will hereafter be designated as the 40–60, 60–60, and 80–60 groups.

Each of the three groups in the radius experiment practiced tracking targets rotating at one of three radii on Day 1: 2.0, 3.5, or 5.0 in. On Day 2 all *S*s in this experiment practiced tracking a target rotating on a 3.5-in. orbit. On both days the speed of rotation was held constant at 60 rpm. These groups will hereafter be designated as the 2.0–3.5, 3.5–3.5, and 5.0–3.5 groups.

Instructions.—All *S*s were given general instructions about the task, suggestions on efficient methods of tracking, and a demonstration. They were told specifically to try to remain on target as long as possible, that they would receive 30 trials lasting for 30 sec. each and that there would be a 60-sec. rest between trials, that they should pick up the handle on the command of "Ready" but not start tracking until the "Start" signal had been given, and that the scoring apparatus would record the number of times they hit the target and how long they stayed on each time. No mention was made of a change in conditions before the start of the second session on Day 2.

Results

Speed of Rotation

Time on target.—Figure 1 presents the mean percentage time on target for both sessions for Groups 40–60, 60–60, and 80–60. As may be seen, when the three groups tracked targets of different speeds, Day 1, their performance was markedly affected. Although a usually effective[2] transformation was applied (4), significant heterogeneity of variance remained ($\chi^2 = 16.20$, $P < .01$, 2 *df*). An extended trend analysis of variance (5) was performed on the transformed Day 1 scores, but because of the heterogeneity of variance the F table was entered with one-half the *df* of the error term. Nonetheless, the means and slopes for the three groups were significantly different, $F = 61.80$ and $F = 8.84$, respectively (2 and 42 *df*

[2] The transformation was essentially logarithmic but since some scores were zero, a constant of .01 was added to all scores. To avoid logs having negative characteristics, another constant of 2.00 was added; therefore the complete transformation was $2 + \log (x + .01)$. This transformation successfully eliminated heterogeneity of variance in all cases except for the time-on-target scores, Day 1, speed experiment.

and $P < .01$ for both). Clearly, time on target performance is acquired more rapidly and to a higher degree as an inverse function of target speed.

On Day 2 all groups practiced under the same conditions of 60 rpm, but their performance curves were markedly dissimilar in certain respects.

An extended trend analysis on the modified log transformation of these scores indicated that the rate of gain as well as the change in rate of gain were highly significant (between group slopes, $F = 6.683$; between group quadratics, $F = 5.766$; 2 and 42 *df* and $P < .01$ for both). The third and fourth order polynomials were also significant but their meaning is unclear.

Table 1 presents the mean percentage of time on target for blocks of 5 trials for each group. The first three columns summarize the data from the speed groups. The remaining three columns summarize the data from the radius groups which will be considered later.

Although the differences among the over-all means for the three groups for Day 2 were not significant in this analysis, there is considerable evidence of differential transfer as measured by level of performance as a function of target speed on Day 1. An analysis of variance of the transformed scores for the first trial for Day 2 shows that the group means differ significantly ($F = 11.60$, $P < .01$, 2 and 42 *df*). A trial-by-trial analysis indicated that the means differed beyond the .05 level on Trials 2, 4, 5, 6, 7, 11, and 12. After Trial 12 no significant differences were obtained. In view of these results

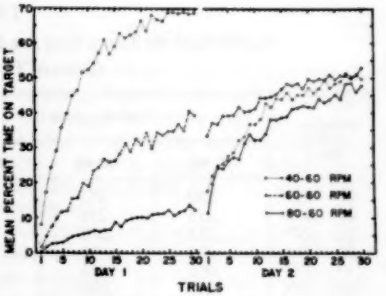

FIG. 1. Mean percentage of time on target for each group in speed experiment on each trial on Day 1 and Day 2.

TABLE 1

MEAN PERCENTAGE OF TIME ON TARGET

Trials	Speed Groups			Radius Groups		
	40–60	60–60	80–60	2.0–3.5	3.5–3.5	5.0–3.5
Day 1						
1–5	22.93	7.02	2.03	25.09	10.89	6.29
6–10	47.11	16.32	4.90	45.82	24.66	16.69
11–15	57.70	25.27	6.85	57.99	36.49	24.31
16–20	63.20	30.36	9.37	64.29	46.57	30.32
21–25	66.65	33.40	11.05	70.77	51.28	34.00
26–30	68.83	37.32	12.03	74.69	55.58	37.34
Day 2						
1–5	23.45	37.20	21.15	47.43	55.19	50.96
6–10	33.64	40.78	30.34	55.94	62.02	56.92
11–15	42.33	45.26	36.39	61.69	66.04	63.26
16–20	45.63	48.47	41.23	63.18	68.47	64.84
21–25	47.65	49.88	43.68	67.23	69.78	66.71
26–30	50.26	51.11	47.05	68.00	71.38	69.20

TABLE 2

PROPORTIONS OF TOTAL HITS FOR SPECIFIED INTERVALS FOR EACH GROUP
(Sums of Trials 1–5 on Day 2)

Duration of Hits (Sec.)	Speed Groups			Radius Groups		
	40–60	60–60	80–60	2.0–3.5	3.5–3.5	5.0–3.5
.00– .049	.280	.162	.226	.087	.082	.076
.05– .099	.235	.231	.296	.162	.149	.166
.10– .199	.251	.255	.255	.283	.238	.267
.20– .299	.104	.152	.120	.170	.150	.171
.30– .399	.065	.063	.045	.078	.085	.087
.40– .599	.045	.086	.044	.117	.140	.120
.60– .799	.015	.029	.009	.044	.071	.055
.80– .999	.003	.010	.004	.023	.031	.017
1.00–1.199	.002	.004	.002	.015	.025	.018
1.20–1.599	<.001	.004	0	.012	.016	.014
1.60–1.999	<.001	.002	0	.006	.010	.006
2.00–4.999	0	.001	0	.003	.004	.003
Over 5.00	0	0	0	0	0	0
Total Hits	3341	3785	2950	3522	3516	3666

and those of the trend test it seems reasonable to conclude that practice at different target speeds shows differential transfer for at least several trials; and that with continued practice, these differences decrease and become statistically insignificant.

Distribution of hits.—The distributions of durations of continuous time on target (hits) obtained from the Tracking-time Analyzer were determined for each group in six blocks of five trials each. The data were grouped to increase the stability of the frequency distribution. The obvious differences on Day 1 will not be described since interest is principally in the transfer effect.

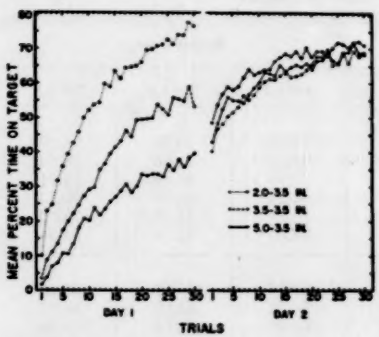

FIG. 2. Mean percentage of time on target for each group in radius experiment on each trial on Day 1 and Day 2.

The left half of Table 2 presents the proportions of these hits at the various durations for the three groups in the speed experiment for Trials 1–5 on Day 2. The 60–60 group not only hit the target more frequently but made significantly more long-duration hits and fewer short-duration hits.

These performance differences in favor of the 60–60 group were consistent throughout all six blocks of five trials (χ^2 values with $P < .001$) although the magnitude of the differences decreased with practice. Evidently the measure of distribution of hits is more sensitive than the usual time-on-target measure.

Target Orbit Radius

Time on target.—The mean times on target for the three groups in this

experiment (2.0–3.5, 3.5–3.5, and 5.0–3.5) for both sessions of practice are presented in Fig. 2. Again, after applying the modified log transformation, an extended trend analysis was performed. As in the speed experiment, the dimension of target orbit radius was highly effective on Day 1. The means differed beyond the .001 level and the between group slopes and quadratics were significant beyond the .05 level.

The data for Day 2 present a different picture of transfer effects than in the speed experiment, however. None of the between-group differences was significant. Therefore, it cannot be concluded that tracking targets of different orbit radii differentially affects performance on a common task, at least as measured by time on target. It may be seen in Table 1, however, that Group 3.5–3.5 consistently scores slightly higher on Day 2 than do the other radius groups. This is in line with the results from Day 2 for the speed groups.

Distribution of hits.—As before, the durations of continuous time-on-target scores were grouped into six blocks of five trials each. These data for the first block of trials (1–5) on Day 2 are shown in the right half of Table 2. As in the speed experiment, the group which did not have a change in conditions (Group 3.5–3.5) displayed a significantly superior level of performance on Day 2. Significant χ^2's were found for all six blocks of Day 2 practice. The *S*s in Group 3.5–3.5 made more long-duration hits than *S*s in either of the other groups.

Once again the distribution-of-hit analyses appeared to be more sensitive in detecting group differences than an analysis of the more usual measure of time on target.

Discussion

A possible reason for different transfer effects as a function of the two methods of varying difficulty is that the levels of difficulty used in the speed experiment gave a greater range of performance than did the levels used in the radius experiment.

An alternative explanation seems more likely, however. If the two major stimulus variables are redefined in more general terms, one can identify variations in speed and radius as comparable to changes in the frequency and amplitude of the stimulus, respectively. Furthermore, when *S* matches his responding rate to the frequency of the stimulus, he will have established a *rhythm* of responding. The greater differential transfer in the speed experiment suggests that this rhythm may be one of the specific component skills which is learned by *S*. Since there was much less differential transfer in the radius experiment (detected only by the continuous hits analyses), the results suggest that the amplitude of response is a much less specific component skill.

It is to be noted that frequency has been equated to speed of rotation and not to target velocity as such, since a change in speed of rotation (frequency) or target orbit radius (amplitude), with the other held constant, will vary target velocity. In the present study, variations in target velocity were actually greater in the radius experiment than in the speed experiment but greater differences in transfer were observed in the latter.

It appears reasonable that at different levels of task difficulty, some component skills will be learned more readily than others. Which skills are learned is probably determined by the ease of practicing the skill, the frequency of reward for such practice, and the ability level of *S*. For example, an *S* practicing with a 5-in. target orbit radius will probably be "off target" more often than an *S* practicing with a 2-in. target orbit radius. The former condition is likely to give *S* more practice in making cor-

rective movements to get back on target. Conceivably the component skill could be a decision-making one. For a high speed of rotation *S* might reasonably decide to wait for the target to catch up to him rather than try to overtake it when he is off target. Such a procedure might be inappropriate when transferring to a slower target rotation, however, since more time would be lost waiting for the target to complete its rotation. The specific decision will of course be determined by both the speed of rotation and the size and direction of the tracking error. It is proposed, therefore, that different methods of varying task difficulty will emphasize the learning of different component skills. Also, a gross time-on-target score will not be an adequate index of the differential component-skill ability.

The persistence on the second task of some component skills is likely because *S* receives *some* over-all reinforcement and may not be able to identify and compare the component skills as such. Such an interpretation would suggest that at various stages of practice of the second task, it should be possible to identify systematic changes in components as prior ones are discarded or modified.

It is conceivable that certain component skills could be emphasized by experimental manipulation of instructions, perceptual feedback, and task requirements, thereby controlling the amount and degree of transfer.

Summary

Two pursuit-rotor experiments were run simultaneously to investigate the effects of varying first task difficulty on transfer to a second task. Two methods of varying difficulty were used: speed of rotation of the target and radius of the target orbit. In the speed experiment, three groups practiced with different speeds on Day 1 (40, 60, and 80 rpm); each group transferred to a common speed (60 rpm) on Day 2. For these groups the target orbit radius was held constant at 5.0 in. In the radius experiment, three groups practiced with different target orbit radii on Day 1 (2.0, 3.5, and 5.0 in.) and transferred to a common radius on Day 2 (3.5 in.). For these groups the speed of target rotation was held constant at 60 rpm. Ninety women *S*s received 30 trials during each of two practice sessions, and each trial lasted 30 sec. with 60-sec. interpolated rests.

The response measures were total time on target and distribution of time continuously on target. The main findings were: (*a*) Either method of varying difficulty of the task produced great variations in Day 1 performance, with decreases in speed or in target orbit radius associated with increased performance. (*b*) The speed variable produced significant differential transfer effects on Day 2 as shown by both response measures. The groups which practiced at 60 rpm on both days performed better than either of the other groups. Practice at 40 rpm produced slightly greater transfer to 60 rpm than did practice at 80 rpm. (*c*) Variation in the radius of the target orbit did not produce significant differential transfer effects on Day 2 as measured by time on target. However, the group which practiced at the same target orbit radius on both days consistently scored a greater proportion of long-duration hits than either of the other groups on Day 2.

References

1. Ammons, R. B., Ammons, C. H., & Morgan, R. I. Transfer of training in a simple motor skill along the speed dimension. WADC Tech. Rep. No. 53–498, 1954.
2. Bourne, L. E., Jr., & Archer, E. J. Time continuously on target as a function of distribution of practice. *J. exp. Psychol.*, 1956, **51**, 25–33.
3. Day, R. H. Relative task difficulty and transfer of training in skilled performance. *Psychol. Bull.*, 1956, **53**, 160–168.
4. Edwards, A. L. Experimental design in psychological research. New York: Rinehart, 1950.
5. Grant, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.
6. Helmick, J. S. Pursuit learning as affected by size of target and speed of rotation. *J. exp. Psychol.*, 1951, **41**, 121–125.

(Received January 2, 1958)

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

EFFECT OF DISTRIBUTION OF PRACTICE ON A COMPONENT SKILL OF ROTARY PURSUIT TRACKING[1]

E. JAMES ARCHER

University of Wisconsin

Both of two current theories of motor learning (**1, 5**) imply that although the *level of performance* under massed practice may be lower than that for distributed practice, the actual *level of learning* is the same for the two conditions. Such a conclusion may be drawn from the following combination of the two theories. Practice of a motor skill results in the acquisition of certain positive habits but simultaneously with the acquisition of these habits there develops a negative (suppressing?) attribute called reactive inhibition (I_R) which, through an unspecified mechanism, results in a temporary work decrement (D_{W_t}). If *S* is allowed to rest periodically—e.g., between trials—I_R is believed to dissipate spontaneously. If, however, *S* is not permitted to rest, it is believed that I_R will build up to a critical level so that when *S* does rest he will learn a negative habit of not responding. This not responding has been called conditioned inhibition ($_sI_R$) and is assumed to be a cause of the relatively poor performance of a prerest massed practice group after an I_R-dissipating rest has been introduced, i.e., permanent work decrement (D_{W_p}). This last conclusion is the least acceptable of the several made by these two theories. With two inhibitory constructs almost any data can be explained since one construct or the interaction of the two will provide "refuge" when the other is inappropriate.

Whereas the foregoing combined theory emphasizes the performance-suppressing character of I_R without suggesting the mode of this inhibitory action, the present paper proposes that the mechanism of performance suppression is *interference* with the practice of the intended skill. An advantage of this latter position is that the construct of $_sI_R$ and its consequent D_{W_p} can be dispensed with thereby making the model more parsimonious.

The results of a previous study (**3**) suggested that since *S*s serving under massed practice show less time on target continuously, a failure to learn, or at least perform the component skill of moving the stylus in a circle, was indicated. Furthermore, when these *S*s showed improvement following a 5-min. rest, they did so by remaining on target for longer durations continuously than they did before the rest. It was hypothesized that this improvement in performance was due to the dissipation of I_R which had been interfering with *S*'s practice of smooth circular movements.

The present study was designed to measure one of the several component skills which is involved in rotary-pursuit tracking: moving the tracking stylus tip in a circle. This com-

[1] Most of the data for this study were gathered by Lyle E. Bourne, Jr. The study was supported by a grant from the Research Committee of the Graduate School from funds from the Wisconsin Alumni Research Foundation. The writer is indebted to the Numerical Analysis Laboratory for its assistance in the data analysis and especially to its Project Supervisor, Eugene L. Albright.

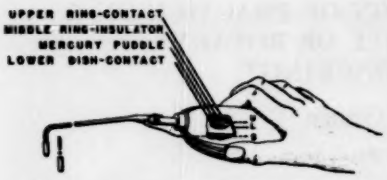

FIG. 1. A schematic diagram of the sensing device which detected noncircular movements and which was imbedded in the handle of the tracking stylus. An exploded view of the spring-loaded ball tip is also shown. The dotted lines U, L, and S indicate three wires which were connected to a Jones plug mounted at the end of the handle. The wire marked S carried the signal from the stylus tip and provided a measure of time on target. The circuit between U and L would be completed if the mercury puddle made contact with the upper ring and the lower dish contacts. This latter signal indicated a circular movement.

ponent skill was selected for study because of its probable sensitivity to interference and the ease of instrumenting its measurement.

Although it has been known for several years (4) that men are superior to women on the rotary pursuit task, little further research has been directed toward analyzing the bases of this difference. On the chance that one reason for this performance difference was a differential ability to make smooth circular movements as well as being differentially resistant to the interference of I_R, an equal number of men and women served in this experiment thereby permitting an analysis of the effect of this variable.

METHOD

Subjects.—Forty-five men and 45 women students in introductory psychology courses at the University of Wisconsin served individually as *Ss*. Fifteen *Ss* of each sex were assigned randomly to one of the three distribution-of-practice conditions. All were naive to the task.

Apparatus and task.—The task was provided by a pursuit rotor. The target was a polished brass disk of .75-in. diameter, set flush with the surface of the turntable. The *S* tracked the target with a stylus which was hinged so as to prevent him from pressing down on the target. As in previous studies in this series (2, 3), the Tracking-time Analyzer was used. In the present study, however, the clock and the counters recorded different response measures. Whenever *S* made contact with the target, the clock was activated. This provided the usual measure of time on target.

A refinement which controlled for the possible change in size and shape of the stylus tip because of wear is shown in Fig. 1. The exploded view on the left shows the spring-loaded steel ball (1/16th-in. diameter) which was the scoring contact of the stylus tip.

A mercury switch of special design was imbedded within the handle of the tracking stylus. A cut-away view of this switch and its position in the handle is also shown in Fig. 1. The three essential parts of the switch are (*a*) an upper ring contact, (*b*) a lower dish contact, and (*c*) a puddle of mercury. The two contacts were ¼ in. apart and insulated from each other. Whenever *S* moved the stylus in a circle the resultant centrifugal force caused the mercury to move out from the depression in the dish contact and complete a circuit with the upper ring contact. Each such circuit completion closed a relay which in turn interrupted the signal to the Tracking-time Analyzer. In this manner the Tracking-time Analyzer was used to obtain a frequency distribution of durations of continuous noncircular movements.

The rotor was operated at 80 rpm instead of the more customary 60 rpm so as to optimize the sensitivity of the mercury switch. Although *Ss* were instructed to hold the stylus handle level, they tended to tilt the handle as if to press down on the target. If the turntable turned at 60 rpm this inclination of the handle was sufficient to cause the mercury to break contact even though *S* was still on target and obviously making a circular movement. It was found that at 80 rpm the centrifugal force was sufficient so that such tilting of the stylus handle did not affect the recording of noncircular movements. The speed of rotation of the rotor was always ascertained by an attached stroboscope before the start of each experimental session.[2]

[2] This method of recording noncircular responses is admittedly somewhat crude. The mercury-puddle device provides a rough dichotomization of movements that are "sort of circular" and other movements. This study was primarily concerned with the other movements. Since *S* could and did rotate and tilt the handle while tracking the target, it was very difficult to get a meaningful calibration of the device. Because the sensing device was not mounted in

In all conditions of practice the turntable revolved continuously throughout the session. The *E* timed the length of the trials and gave "ready," "start," and "stop" signals orally to *S*. The plate circuit of the Tracking-time Analyzer was turned *on* only during the trials, thereby preventing the scoring of noncircular movements except during a trial.

The operation of the analyzer was clearly audible to *S* since it was separated from him only by a partial wall. The calibration of the analyzer was checked at least twice a day to correct for changes in line voltage.

Conditions.—All *S*s received 21 30-sec. trials with either 0-, 30-, or 60-sec. rest between each trial. After Trial 21 all *S*s had a 10-min. rest, which was then followed by nine trials with no intertrial rests. During all rest periods *S* was encouraged to read magazines which were on a table behind him when he faced the pursuit rotor. This activity kept *S*s from visually tracking the target during the rest periods.

Results

Prerest Performance

First trial measures.—The performances of the six groups on the *first* trial were compared in terms of four different response measures: time on target, time of noncircular movements, frequency of noncircular movements, and frequency distribution of noncircular movement durations. For none of these measures was the intertrial rest interval statistically significant (since none of the *S*s had had a rest, any differences on this variable could only be due to chance). For all measures except time on target there was a sex difference on the first trial. Men spent less time making noncircular movements than women, 3.30 sec. vs. 5.58 sec. ($F = 13.18$, 1 and 84 *df*, $P < .001$). Men made fewer noncircular movements than women, 34.02 vs. 44.06, ($F = 6.28$, 1 and 84 *df*, $P < .025$). Furthermore, men made fewer long-duration and more short-duration noncircular movements ($\chi^2 = 50.81$, 6 *df*, $P < .001$).

Time on target.—The mean percentage of time on target for each of the six groups for all 30 trials is shown in the lower half of Fig. 2. A summary of an extended trend analysis of variance is presented in Table 1. The assumption of homogeneity of variance was met after applying a modified logarithmic transformation; $2 + \log (\text{TOT} + .01)$.

This trend analysis indicates that although the sexes differ in terms of mean time on target (Line 1), neither their slopes (Line 5) nor change in slopes (Line 9) differ significantly. On the other hand, the means (Line 2), slopes (Line 6), and change in slopes (Line 10) increase as the intertrial rest interval is increased.

Frequency of noncircular movements.—As expected, the frequency of noncircular movements was inversely related to the duration of the intertrial rest interval: 31.85, 22.27, and 19.33 for 0-, 30-, and 60-sec. rest, respectively ($F = 11.52$, $P < .01$ for 2 and 84 *df*). Furthermore, men made fewer noncircular responses than women, 21.16 vs. 27.80 ($F = 8.464$, $P < .01$, for 1 and 84 *df*). Finally, during practice the frequency of noncircular movements decreased at a negatively accelerated rate from an average of 39.04 on the first trial to 21.44 on the last. Also the frequency distribution of these noncircular movements changed from near normal on the first trial to markedly skewed later in practice.

Table 2 presents the frequency

gimbals, the orientation of the handle in other than the horizontal plane adversely affected the sensitivity of the device. Although it would have provided a more direct method of recording the noncircular responses, a motion picture recording system was not used because of the difficulty of quantitative evaluation of such records.

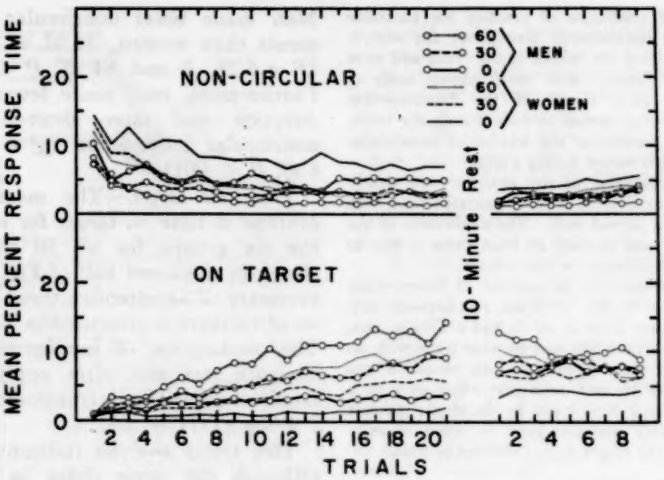

FIG. 2. The mean percentage of response time for both noncircular movement and time on target is shown as a function of the three intertrial rest intervals and the two sexes.

distributions for some of the prerest and all of the postrest trials. The prerest data (Trials 1–21) give but a crude estimate of the changes since the distributions for different sexes and practice schedules have been combined to obtain more stable values. The general effect of practice is evident, however. The number of *Ss* making many noncircular movements decreases and the number making few such movements increases.

TABLE 1

SUMMARY OF EXTENDED TREND ANALYSIS OF VARIANCE OF PREREST TIME ON TARGET SCORES; 2 + LOG (TOT + .01)

Line	Source	*df*	*MS*	Error	*F*
1	Between sex means	1	27.973	4	12.595**
2	Between rest means	2	63.812	4	28.731**
3	Between sex × rest means	2	0.441	4	
4	Between individual means	84	2.221	17	20.009**
5	Between sex slopes	1	0.482	8	1.377
6	Between rest slopes	2	6.858	8	19.594**
7	Between sex × rest slopes	2	0.322	8	
8	Between individual slopes	84	0.350	17	3.153**
9	Between sex quadratics	1	0.665	12	2.854
10	Between rest quadratics	2	2.442	12	10.481**
11	Between sex × rest quadratics	2	0.178	12	
12	Between individual quadratics	84	0.233	17	2.099**
13	Between groups residual	90	1.061	17	9.559**
14	Over-all slope	1	122.628	8	350.366**
15	Over-all quadratic	1	7.386	12	31.700**
16	Over-all residual	18	0.292	17	2.631*
17	Between individual residual	1512	0.111		
18	Total	1889			

* $P < .01$.
** $P < .001$.

TABLE 2

NUMBER OF *Ss* HAVING DIFFERENT FREQUENCIES OF NONCIRCULAR MOVEMENTS FOR VARIOUS TRIALS

Trial	Number of Noncircular Movements									r $N = 90$
	0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	80–89	
Prerest										
1	6	8	18	19	15	10	8	1	5	−.42
5	14	20	16	22	11	5	2			−.46
9	21	21	19	17	7	5				−.39
13	16	31	18	16	7	2				−.53
17	22	26	16	18	5	3				−.56
21	21	24	20	17	5	2		1		−.44
Postrest										
22	43	25	10	8	4					−.18
23	36	25	15	7	6	1				−.16
24	31	26	16	9	4	4				−.26
25	29	26	18	12	5					−.18
26	23	33	15	12	6	1				−.25
27	25	25	20	13	5	2				−.27
28	24	27	14	20	3	2				−.29
29	22	20	21	17	8	1	1			−.34
30	28	21	13	17	7	1	2	1		−.31

Time of noncircular movement.—The mean percentage of time of noncircular movement for each of the six groups is also shown in Fig. 2. A modified logarithmic transformation was applied before these scores were subjected to an extended trend analysis of variance. A summary of this analysis is presented in Table 3. Because none of the between-group terms beyond the linear component was significant, the table has been condensed to save space. As in the time-on-target trend analysis of variance there was a significant sex difference; men spent less time making noncircular movements than did women (Line 1). Furthermore, amount

TABLE 3

SUMMARY OF EXTENDED TREND ANALYSIS OF VARIANCE OF PREREST TIME OF NONCIRCULAR MOVEMENTS; 2 + LOG (TNR + .01)

Line	Source	*df*	*MS*	Error	*F*
1	Between sex means	1	25.686	4	12.122*
2	Between rest means	2	24.296	4	11.466*
3	Between sex × rest means	2	2.476	4	1.168
4	Between individual means	84	2.119	12	36.534*
5	Between sex slopes	1	0.525	8	1.836
6	Between rest slopes	2	2.664	8	9.315*
7	Between sex × rest slopes	2	0.487	8	1.703
8	Between individual slopes	84	0.286	12	4.931*
9	Between group residual	95	0.062	12	1.069
10	Over-all slope	1	29.026	8	101.490*
11	Over-all residual	19	0.478	12	8.241*
12	Between individual residual	1596	0.058		
13	Total	1889			

* $P < .001$.

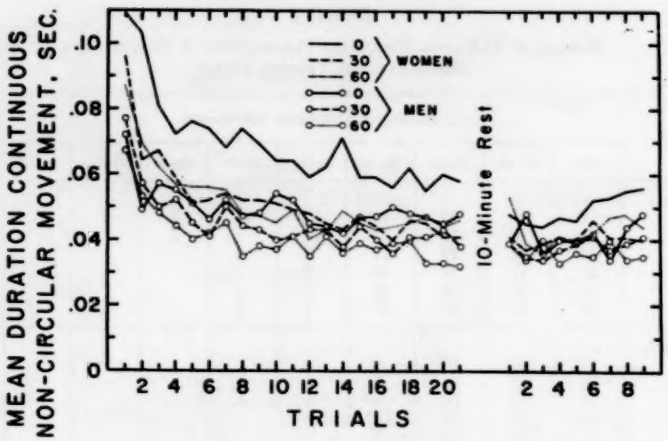

FIG. 3. Mean duration of continuous noncircular movement in seconds for all distribution-sex combinations for all 30 trials.

of noncircular movement time was inversely related to the intertrial rest interval (Line 2). Of particular interest is the observation that with practice the amount of time spent in noncircular movement decreases more rapidly and attains a lower asymptote as the intertrial rest interval increases (Line 6).

Derived noncircular movement measures.—The mean duration of noncircular movement was computed for each *S* for each trial. These values were determined by dividing the total time of noncircular movement per trial by the number of such movements. The results, which are presented in Fig. 3, indicate that the effect of distribution of practice and sex difference which was shown in the upper part of Fig. 2 was not simply due to the increased frequency of noncircular movements. Not only do women make *more* noncircular movements than men, but the average duration of each movement is longer (for all prerest trials, $F = 21.068$, 1 and 84 *df*, $P < .001$). A similar conclusion may be drawn from the effect of distribution of practice: not only does decreasing the intertrial rest interval increase the frequency of noncircular movements but it also increases the average duration of the noncircular movement response (for all prerest trials, $F = 8.172$, 2 and 84 *df*, $P < .001$). Furthermore, the rate of improvement with practice for women is probably greater than that for men (between-sexes linear term was significant, $F = 4.685$, 1 and 84 *df*, $P < .05$). None of the other between-groups terms was significant.

In order to test for a resting response (6) the *longest* duration of continuous noncircular movement was determined for each *S* for each trial. As with each of the previous measures a modified log transformation was applied to eliminate heterogeneity of variance before performing an extended trend analysis of variance. Except for individual and over-all significant deviations, only the between-sex means and between-rest-

interval means were significant sources of variation ($F = 15.696$, 1 and 84 *df*, and $F = 11.020$, 2 and 84 *df*, respectively). None of the higher order terms was significant. The absence of a significant between-slopes term argues against the hypothesis that *S*s, under massed practice, will learn the habit of not responding—i.e., conditioned inhibition.

Postrest Performance

Time on target.—As expected, the greatest increase in performance following the rest occurred in the prerest group which had had a 0-sec. intertrial rest. The smallest increase was for the 30-sec. prerest group. The group which had had a prerest 60-sec. intertrial interval showed a decrease in performance after the rest. These differences were significant beyond the .05 level ($F = 4.092$, 2 and 84 *df*). Sex of *S* did not differentially affect the amount of recovery from I_R.

Not only did the prerest practice conditions affect the amount of change in performance to the first postrest trial, it also affected the level of this performance. As may be seen from the postrest performance curves in the lower half of Fig. 2, the level of performance on the first postrest trial was directly related to the duration of the prerest intertrial rest interval. These differences were significant beyond the .001 level ($F = 9.817$, 2 and 84 *df*), and are usually taken as an index of sI_R. The performance of the two sexes was probably still different also ($F = 4.665$, 1 and 84 *df*, $P < .05$).

The log transformations of postrest time-on-target scores were analyzed with an extended trend analysis of variance. Only two sources of variance were significant: between means and between slopes for the different intertrial rest intervals ($F = 6.896$, 2 and 84 *df*, $P < .001$) and ($F = 3.653$, 1 and 84 *df*, $P < .05$), respectively. These results indicate that the apparent convergence of the postrest performance curves shown in the lower half of Fig. 2 was significant and that over all of the postrest trials the level of performance was proportionate to the prerest degree of distributed practice.

Frequency of noncircular movements.—As expected, the frequency of noncircular responses decreased as an inverse function of the prerest intertrial rest interval ($F = 8.411$, $P < .01$ for 2 and 84 *df*) from the last prerest-trial to the first postrest trial. This reduction was unrelated to the sex of *S*. Although the prerest practice condition affected the magnitude of *decrease* in frequency of noncircular movements, it did not affect the final level of responding. Under comparable conditions (massed practice for all groups) the frequency of noncircular movement was unrelated to prerest intertrial conditions.

The effect of massed practice (the uniform postrest condition) is evident from the lower half of Table 2 (Trials 22–30). As practice continues, and I_R presumably accumulates, the number of noncircular movements increases as well as the interindividual variation.

Derived noncircular movement measures.—The total time of noncircular movement was unrelated to the prerest conditions. The sex difference remained, however ($F = 7.576$, $P < .01$ for 1 and 84 *df*).

In summary, except for this sex difference none of the prerest effects was significant in the analyses of postrest scores. Men made fewer noncircular responses which were on

TABLE 4

INTERCORRELATIONS OF MEASURES OF NONCIRCULAR MOVEMENTS (EXCEPT FREQUENCY) AND TIME ON TARGET

Variable	Prerest (N = 1890)			Postrest (N = 810)		
	Time, Noncircular	Mean Duration Noncircular	Longest Duration Noncircular	Time, Noncircular	Mean Duration Noncircular	Longest Duration Noncircular
Time on target	−.44	−.36	−.24	−.28	−.25	−.24
Time, noncircular		.77	.57		.75	.72
Mean duration, noncircular			.78			.78

the average of shorter duration and, over all, the time of noncircular movement was less for men. Furthermore the mean longest duration of noncircular movement was also less for men.

Correlations of Measures

Although the measures of frequency and duration of noncircular movements became increasingly more skewed with practice (see Table 2), it was of some interest to know of the correlations of these measures and the usual performance measure of time on target. The intercorrelations for all of these measures, except frequency of noncircular movements, were determined for each group for each trial both pre- and postrest. The only detectable trends as a function of practice were (*a*) a marked drop in the correlations immediately after the 10-min. rest and (*b*) a gradual increase in the correlations during the postrest trials as shown in the last column of Table 2. The over-all pre- and postrest correlations between frequency of noncircular movements and time on target were −.48 and −.25, respectively. The other intercorrelations for the less severely skewed measures are presented in Table 4.

DISCUSSION

The results of the present study lend considerable support to the following expansion of a miniature behavioral system. During and because of the practice of a motor skill certain negative factors (I_R) accumulate. If rests of sufficient frequency and duration are interpolated between practice sessions, these negative factors will dissipate without affecting the level of performance. If, however, the rests are of insufficient duration, the negative factors will accumulate to a sufficient level to cause an involuntary tremor of the effector system practicing the skill. In a perceptual-motor skill such as used in this study, the effect of these tremors produced and were recorded as a failure to make circular movements.

Since the acquisition of the rotary pursuit skill requires the learning of the circular movement response, it would be expected that the *rate* of such learning would be affected by the length of the intertrial rest interval. As the intertrial rest interval decreases, the opportunity for the accumulation of I_R increases and its observable effect—tremor-produced noncircular movements—increases. Whereas the *learning* of the rotary pursuit skill decreases the frequency and duration of noncircular movements, the massing of practice operates as a counter effect; hence different rates of learning. That this did happen is evident from the significant

between-rest interval slopes term in Table 3 and from Fig. 2.

Furthermore, we would expect that one effect of a long interpolated rest for all groups would be to markedly reduce the frequency and total duration of noncircular movements since, with rest, the negative factor which produced the tremors, which in turn produced noncircular responses, will spontaneously dissipate. This also was observed in the present study.

Finally, when practice is resumed there is no reason to expect that the groups which had had differing prerest intertrial rests will perform equally well. The failure to perform equally well has previously been regarded as evidence for conditioned inhibition or permanent work decrement by some writers (**1, 5**). However, when the practice-interfering effect of the tremor response is considered, it would be more remarkable if the groups did *not* differ on the first postrest practice trial. After all, they have not had an equal opportunity to practice, e.g., a smooth circular movement and a tremor are incompatible. In the present study, the level of performance (as measured by time on target) on the first postrest trial was significantly affected by the prerest intertrial interval; the level was proportional to the intertrial rest interval. Since the amount of interference with the practice of the rotary pursuit skill was inversely related to the intertrial rest interval, it is proposed that the concepts of ${}_{s}I_{R}$ and D_{W_r} are no longer necessary since differential practice is a more parsimonious explanation. Also, irrespective of the prerest intertrial conditions, the frequency and total duration of noncircular movements increased when the postrest practice was massed. In short, the degree to which I_R interferes with the practice of a skill appears to be a simple inverse function of the distribution of that practice.

At least a partial explanation of the time-on-target differences which are observed for the two sexes can be found in the present study. For naive *S*s, men make fewer noncircular responses and, on the average, these are of shorter duration than comparable measures for women. Although both sexes learn to reduce the frequency and duration of these inappropriate tracking responses, men are still superior in this component skill. Neither sex, however, was better at resisting the effects of massing of practice.

Summary

In a rotary pursuit experiment, the frequency and duration of noncircular movements was measured as a function of intertrial rest interval and sex of the *S*. As expected, the frequency and duration of noncircular movements was inversely related to intertrial interval. Also, men made fewer such movements and they were shorter than for women. With practice the number and duration of noncircular movements decreased. Furthermore, the rate of decrease was directly related to the intertrial interval.

After a 10-min. rest the frequency and duration of noncircular movements decreased as an inverse function of the prerest intertrial rest interval. The prerest interval did not affect the frequency and duration of noncircular movements although it did directly affect the duration of time on target.

The following hypotheses were proposed: (*a*) The reactive inhibition (I_R), which develops during and because of practice of a motor skill, will produce tremors. These involuntary reactions prevented *S* from making smooth circular movements. (*b*) Performance, as measured by time on target, is poorer as the intertrial rest interval is decreased because the tremors interfere with the practice of the rotary pursuit skill. (*c*) Because of this interference, the skill is learned to a lower degree and this differential practice is a simpler explanation of the apparent permanent work decrement than is the concept of conditioned inhibition.

References

1. Ammons, R. B. Acquisition of motor skills: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, **54**, 263–281.
2. Archer, E. J., & Bourne, L. E., Jr. Inverted-alphabet printing as a function of intertrial rest and sex. *J. exp. Psychol.*, 1956, **52**, 322–328.

3. Bourne, L. E., Jr., & Archer, E. J. Time continuously on target as a function of distribution of practice. *J. exp. Psychol.*, 1956, **51**, 25–33.
4. Buxton, C. E., & Grant, D. A. Retroaction and gains in motor learning: II. Sex differences, and a further analysis of gains. *J. exp. Psychol.*, 1939, **25**, 198–208.
5. Kimble, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, **39**, 15–23.
6. Kimble, G. A. Performance and reminiscence in motor learning as a function of the degree of distribution of practice. *J. exp. Psychol.*, 1949, **39**, 500–510.

(Received January 13, 1958)

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

EFFECT OF STIMULI TIME RELATIONS DURING PRECONDITIONING TRAINING UPON THE MAGNITUDE OF SENSORY PRECONDITIONING[1]

DONALD R. HOFFELD,[2] RICHARD F. THOMPSON,[3] AND W. J. BROGDEN
University of Wisconsin

Since the report of the initial study of sensory preconditioning by Brogden (**1**), studies have been reported that confirm the phenomenon (**2, 5, 9**), but those that have attempted to show variation in the magnitude of sensory preconditioning as a function of several independent variables have been unsuccessful (**5**). The lack of success may be due to the fact that these studies were conducted with human beings as Ss. It appears then that experimental studies with animals may be more productive in identifying variables that affect the magnitude and occurrence of sensory preconditioning phenomena.

The present study was designed to investigate the relation between the time of onset and duration of the stimuli involved in preconditioning training upon the magnitude of sensory preconditioning in cats. Since Wolfle (**12**), Kappauf and Schlosberg (**8**), and Spooner and Kellogg (**11**) have reported variation in amount of conditioning as a function of the time relations of the CS and UCS during training, the magnitude of sensory preconditioning should be a function of the time relations of the stimuli involved during preconditioning. If the one of these two stimuli to be made a CS later by avoidance training is designated the UCS during preconditioning, then the other (the one involved in the critical test following both preconditioning and avoidance conditioning training) may be designated the CS. This framework then provides the opportunity to vary the time relations of the CS and UCS during preconditioning training. Silver and Meyer (**10**) compared the relative effectiveness of backward, simultaneous, and forward time relations during preconditioning upon sensory preconditioning measured by transfer conditioning training. They obtained evidence of a greater amount of sensory preconditioning for the forward preconditioning than for backward and simultaneous preconditioning. There was no difference between backward and simultaneous preconditioning. The present study is concerned with the effect of the magnitude of precedence of the CS over the UCS (as defined above) during preconditioning upon the magnitude of sensory preconditioning. It was designed to test the hypothesis that the magnitude of sensory preconditioning will increase with an increase in precedence of the CS over the UCS from zero up to an optimum value and thereafter decline with further increase in the precedence of CS over UCS.

[1] Supported in part by the Research Committee of the Graduate School from funds granted by the Wisconsin Alumni Research Foundation.

[2] Predoctoral Fellow, National Science Foundation.

[3] Postdoctoral Fellow, National Institutes of Health.

PROCEDURE

The experiment was conducted in a sound-insulated, light-shielded room. The condition-

ing apparatus was the rotating cage of Brogden and Culler (4). The sensory conditioning stimuli were a tone and a light. The tone was 1000 cycles in frequency and was 60 db above .0002 dyne/cm.[2] The sound source was a loudspeaker in the experimental room located in a corner 1 ft. in front of the rotator. The light stimulus was provided by the illumination of two 100-w. light bulbs. The light bulbs were hung from the ceiling so that one bulb was 4 in. in front of the cage and the other bulb was 4 in. from the back of the cage. Thus, when the light stimulus was presented, *S* always faced an illuminated bulb. The *E* made his observations through a one-way glass in the wall of the experimental room. Apparatus for the production of the stimuli and their time control was located outside of the experimental room. Tone and light durations were controlled by Tektronic Model 162-A wave form generators. The tone was fed to an electronic switch with a rise-fall time of 50 msec. Electric shock was produced by an electronic stimulator that automatically maintained a constant current. During avoidance training the intensity of the shock was adjusted by *E* to the minimum necessary to provide a minimum cage-turning UCR. The duration of shock was .1 sec. and it occurred at the end of the light CS.

The 24 cats of the experiment were assigned at random so that *N* for each of the five experimental and one control groups was four. Two *E*'s tested *S*s so that each *S* was conditioned and tested by a single *E* but *E* was not orthogonal in the experimental design. The daily experimental session involved 20 trials of the appropriate condition. For each group there were four procedures: (*a*) 10 test periods of preconditioning training; (*b*) instrumental avoidance training with a 2-sec. light CS and .1-sec. shock UCS until the conditioning criterion of 90% CR (18 or more cage-turning shock-avoiding CR's for the 20 trials was reached); (*c*) tests for sensory preconditioning by presenting a 2-sec. tone CS until frequency of CR in a given test period was 10% or less; (*d*) extinction of the CR to the light CS until the criterion of 10% or less CR in a given test period was reached. During preconditioning training the light was always of 2-sec. duration. Preconditioning for the experimental groups involved pairing of tone (CS) and light (UCS). The tone always terminated when the light terminated but it preceded the onset of the light by 0 sec. (E_1), .5 sec. (E_2), 1.2 sec. (E_3), 2 sec. (E_4), and 4 sec. (E_5). Thus the duration of the tone for the five experimental groups respectively was 2, 2.5, 3.2, 4, and 6 sec. The *S*s were quiet during preconditioning training. Since no cage-turning response occurred to any presentation of the preconditioning stimuli, this requirement of the sensory preconditioning paradigm (1) was met. The *S*s of the control group (C) were placed in the rotator for 10 min. on each of the 10 days of pretraining, but received no stimulation of tone or light. The conditions of the remaining three procedures were the same for all groups. During instrumental avoidance training a CR was defined as a cage-turning response that occurred before the termination of the CS. The occurrence of a CR provided for avoidance of the electric shock UCS. This same criterion of response was maintained for the third and fourth procedures except that no shock was involved. The intertrial interval for all procedures varied between 15 and 30 sec.

Results

The data on acquisition of the conditioned cage-turning response to the light CS was examined first to determine whether or not there were

TABLE 1

Results of Posttest Trials of Tone

S	Group					
	C	E_1	E_2	E_3	E_4	E_5
1	0 (1)	6 (2)	6 (2)	4 (2)	7 (2)	15 (3)
2	0 (1)	3 (2)	5 (2)	2 (1)	4 (2)	10 (3)
3	0 (1)	2 (1)	2 (1)	4 (2)	7 (2)	6 (2)
4	0 (1)	2 (1)	1 (1)	1 (1)	4 (2)	7 (2)
Mean Responses	0.00	3.25	3.50	2.75	5.50	9.50
Mean Test Periods	1.00	1.50	1.50	1.50	2.00	2.50

Note.—The first number in each cell is the frequency of response to the tone. The second number in parentheses is the number of test periods required to meet the extinction criterion of two or less responses in one test period.

differences between the groups that might affect the measures of sensory conditioning. An analysis of variance of the number of trials to reach the acquisition criterion was performed and no significant differences between groups were obtained. If there is no evidence that preconditioning has a differential effect upon acquisition of the cage-turning CR to the light CS, there can be no differential effect of acquisition training upon the measures of sensory preconditioning.

The test for cross modal stimulus generalization provided by the frequency of response to the trials of tone alone is the standard measure for sensory preconditioning. Since these tests involve extinction of the generalized CR to the tone, any differences in amount of cross modal generalization might be reflected also as extinction transfer effects to the extinction of the CR to the light CS. Examination of the frequency of CR, the number of trials and the number of test periods to the extinction criterion of 10% or less for the CR to the light CS showed no apparent differences. Analyses of variance of these data produce in each case an insignificant F ratio of the between- and within-group variances. Frequency of response to the tone in the test periods immediately following acquisition, however, is different for the control and experimental groups, and there appear to be differences between the experimental groups.

Table 1 presents the frequency of response for each S in each group and the number of test periods required to meet the extinction criterion of two or less responses in one test period. The group means of these two measures are given also. Since none of the *Ss* in the control group gave any response to the tone and all *Ss* in the experimental group gave some responses to the tone, there is clear evidence of sensory preconditioning.[4] An analysis of variance of these data provides a significant F ratio and supports this conclusion. The increase in magnitude of sensory preconditioning particularly at the greater durations of tone during preconditioning training for Groups E_4 and E_5 was examined further by an analysis of variance of the frequency of response to tone during the critical test for the experimental groups only. The results of this analysis are given in Table 2.

TABLE 2

ANALYSIS OF VARIANCE OF RESPONSES TO TONE BY EXPERIMENTAL GROUPS

Source	*df*	*MS*	*F*
Between Groups	4	30.83	5.00*
Linear	1	78.40	12.71*
Quadratic	1	34.57	5.60*
Cubic	1	2.03	0.33
Quartic	1	2.60	0.42
Within Groups	15	6.17	

* $P = .05$.

The orthogonal polynomial analysis of variance (7) provides not only for an F test of the between-groups and within-groups variances but also provides a test of the fit of linear, quadratic, cubic and quartic functions

[4] The intended design of the experiment included a second control group for noncontiguous presentation of tone and light. This group should have been given pretraining trials of the tone alone, but through an error was given pretraining trials of the light alone. One of the four *Ss* of this group gave two responses to the tone alone following conditioning of the cage-turning response to the light. The others gave no responses. There are no significant differences between this group and the standard control group and there are significant differences between this group and all the experimental groups. The failure to find any effect from control groups for noncontiguous presentation of the preconditioning stimuli (2, 5, 10) makes this control procedure unnecessary.

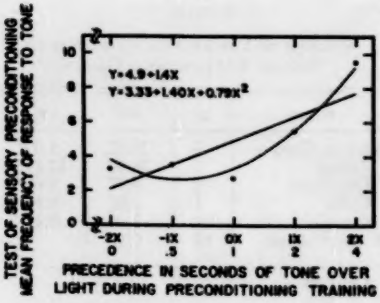

FIG. 1. Relation between magnitude of sensory preconditioning and precedence of tone over light during preconditioning.

to the relation between magnitude of sensory preconditioning and degree of precedence of the tone over the light during preconditioning training. The significance of the between-groups source of variation demonstrates significant differences between the group means. The significance of the linear and quadratic sources demonstrates that there are significant linear and quadratic components in the relation of magnitude of sensory preconditioning and magnitude of precedence of the tone over light and that a quadratic function provides the best fit for these data.

The group means for the five experimental groups are plotted in Fig. 1 and the solid straight and curved lines are the best fitting linear and quadratic functions obtained from the analysis, the equations for which are also given in Fig. 1. A further check on the functional relation of magnitude of sensory preconditioning and precedence of the tone over the light was obtained by the Duncan Range Test (6) of the differences between group means. The results of this test show that the mean of 9.5 responses for Group E_5 is significantly different from the means for each of the other groups and that there are no other significant differences between means.

DISCUSSION

The present data confirm the phenomenon of sensory preconditioning in another species of animal with another type of response and establish a difference in the magnitude of sensory preconditioning as a function of the time relations between the stimuli during preconditioning training. In dealing with the latter result it appears to be unwise to go beyond the results obtained with the Duncan Range Test. This test establishes a significantly greater amount of sensory preconditioning for Group E_5 (precedence of tone over the light by 4 sec. during preconditioning) than that for any of the other experimental groups where the precedence of tone over light was less. Neither the linear nor the quadratic equations obtained from the orthogonal polynomial analysis of the data appear entirely reasonable. Both of these functions suggest that with preconditioning in which the precedence of tone over light is greater than 4 sec. the magnitude of sensory preconditioning will increase.

There must, however, be a limit to this further increase in magnitude of sensory preconditioning since at some value of precedence of the tone over light, conditions for the formation of a delayed CR will be produced, and the conditions for delayed conditioning training are relatively inefficient in the amount of conditioning per unit of training (3). On the other hand, it is possible that the phenomenon of sensory preconditioning is sufficiently different from standard conditioning that the same relations between magnitude of conditioning and time relations between the stimuli do not hold. If one considers what might be expected for time relations during pretraining in which the light UCS precedes the tone CS, i.e., backward conditioning training, the

linear function appears to be more reasonable than the quadratic function. In view of the results obtained by Wolfle (12) and Spooner and Kellogg (11), one would expect a further decline to zero in the magnitude of sensory preconditioning as the precedence of the light UCS over the tone CS is increased. Again it may be argued, however, that the phenomenon of sensory preconditioning is sufficiently different from the standard conditioning so that this relationship might not hold. The failure of Silver and Meyer (10) to find any difference in amount of sensory preconditioning between backward and simultaneous pretraining supports this view.

In any case, the results of the present experiment support without question the conclusion that the magnitude of sensory preconditioning for the conditions of Group E_5 is greater than for the conditions of the other four experimental groups. Whether there is a continuous function of the magnitude of sensory preconditioning and the time relations of the stimuli involved in preconditioning training will require experimental study involving greater degrees of precedence than those of the present experiment and also variation in the time relations of the backward conditioning variety.

Summary

The experiment was designed to investigate the relation between the time of onset and duration of the stimuli involved in preconditioning and the magnitude of sensory preconditioning. During the 200 trials of preconditioning the light (UCS) was 2 sec. in duration and terminated at the same time as the tone (CS). For the five experimental groups, the onset of the tone preceded the onset of the light by 0 sec. (E_1), .5 sec. (E_2), 1.2 sec. (E_3), 2 sec. (E_4), and 4 sec. (E_5). The cats of the control group for cross modal sensory generalization were placed in the rotator for 10 min. a day for 10 days as a control on preconditioning training. All groups received avoidance training until the criterion of 18 or more shock-avoiding cage-turning CR's to the light CS in one test period was reached. Twenty trials of the tone alone were given in daily test periods until the frequency of CR was two or less. Finally, extinction trials of 20 daily presentations of the light were given until frequency of CR was two or less.

Analysis of variance showed no significant differences between groups in speed of acquisition or in any measures of extinction to the light. Significant differences were obtained for frequency of response to the tone. There were no responses to the tone by *S*s of the control group. Some response to the tone was made by every one of the experimental *S*s. An orthogonal polynomial analysis of variance showed significant differences between the experimental groups with significant linear and quadratic components in the relation between magnitude of sensory preconditioning and the degree of precedence of tone over light during pretraining. The Duncan Range Test, however, showed only that the mean for Group E_5 was significantly greater than the means of all other groups. Although there is clear evidence that amount of sensory preconditioning varies as a function of the time relations of stimuli involved during preconditioning, the results of the present experiment do not establish a precise, continuous function. The suggestion is made that the effect of time relations during preconditioning may affect sensory preconditioning in different ways than do time relations of the CS and UCS during standard conditioning.

References

1. Brogden, W. J. Sensory preconditioning. *J. exp. Psychol.*, 1939, **25**, 323–332.
2. Brogden, W. J. Sensory preconditioning of human subjects. *J. exp. Psychol.*, 1947, **37**, 527–539.
3. Brogden, W. J. Animal studies of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
4. Brogden, W. J., & Culler, E. Device for the motor conditioning of small animals. *Science*, 1936, **83**, 269–270.
5. Chernikoff, R., & Brogden, W. J. The effect of instructions upon sensory preconditioning of human subjects. *J. exp. Psychol.*, 1949, **39**, 200–207.
6. Duncan, R. B. A significance test for differences between ranked treatments in an analysis of variance. *Va. J. Sci.*, 1951, **2**, 171–189.
7. Grant, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.
8. Kappauf, W. E., & Schlosberg, H. Conditioned responses in the white rat.

III. Conditioning as a function of the length of the period of delay. *J. genet. Psychol.*, 1937, **50**, 27–45.

9. KARN, H. W. Sensory preconditioning and incidental learning in human subjects. *J. exp. Psychol.*, 1947, **37**, 540–544.
10. SILVER, C., & MEYER, D. R. Temporal factors in sensory preconditioning. *J. comp. physiol. Psychol.*, 1954, **47**, 57–59.
11. SPOONER, A., & KELLOGG, W. N. The backward conditioning curve. *Amer. J. Psychol.*, 1947, **80**, 321–334.
12. WOLFLE, H. M. Conditioning as a function of the interval between the conditioned and the original stimulus. *J. gen. Psychol.*, 1932, **7**, 80–103.

(Received January 13, 1958)

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

A COMPARISON OF CORRECTION AND MODIFIED CORRECTION PROCEDURES ON THE ACQUISITION OF A 12-UNIT VERBAL MAZE [1,2]

RICHARD F. THOMPSON

University of Wisconsin

The experiment presented in this paper compares the effects of two learning procedures on the acquisition of a verbal maze. In previous studies by Brogden and Schmidt (**1, 2**) on this type of maze the correction procedure was employed, a necessary result of maintaining the analogy with spatial maze situations. In the verbal maze, *S* is allowed to make only a limited number of responses, the responses being the alternate choices per maze unit.

The terms "correction" and "modified correction" refer to treatment of *S* in a learning situation. Under the former procedure, *S* must respond until the correct response is discovered. Under the modified correction procedure, more typically employed in serial verbal learning, *S* responds only once and is then presented with the correct response which he may or may not be required to repeat. It will be noted that the number of correct responses presented to *S* in a given number of trials is the same under both procedures, the differences lying in the greater number of responses required from *S* in correction learning.

Under the correction procedure, when number of alternate choices per choice point is increased (**1, 2**), amount of discovery required for a given amount of information increases, and difficulty of acquisition as measured by time and errors also increases. It might be expected, then, that an increase in number of alternate choices would result in a disproportionate increase in difficulty in the correction procedure, relative to the increase in difficulty in the modified correction procedure, where increase in number of alternate choices does not change the amount of discovery required.

Procedure

Five mazes were used, each 12 units in length, with 2, 3, 4, 5, or 6 alternate choices per maze unit. Each unit of Maze 1, the 2-choice maze, included a choice of right (R) or left (L). In Maze 2, the three choices were right, left, and up (U). For Maze 3, the choices were right, left, up, and down (D). The Maze 4 pattern included right, left, up, down, and right-up (RU). Finally, in Maze 5, the six choices were right, left, up, down, right-up, and left-up (LU). The correct patterns for the mazes have been presented elsewhere (**8**).

In the previous studies (**1, 2**), the pattern of correct choices for each maze was chosen from a table of random numbers, the only restriction being that an equal number of each kind of choice be used, except where impossible (e.g., 6 choice, 16-unit maze). However, when using tables of random numbers, it sometimes

[1] Supported in part by a grant from the Research Committee from funds provided by the Wisconsin Alumni Research Foundation

[2] The author wishes to express his indebtedness to W. J. Brogden, whose aid and encouragement made possible this study.

happens that different patterns chosen are unequal in relative difficulty. Hence, in the present study the following additional criteria were employed: (*a*) One symbol appears no more than twice in succession; (*b*) in Mazes 2–5, a doublet cannot appear in immediate conjunction with another such doublet; (*c*) in Mazes 2–5, no doublet is allowed to appear more than twice; (*d*) in Mazes 2–5, no doublet is allowed to follow itself immediately; and (*e*) in Mazes 2, 4, and 5, "up" and combinations with "up" rather than "down" and combinations with "down" are used. In previous studies, in patterns where one or the other could be used, no significant differences were found between patterns containing "up" and combinations with "up" and identical patterns substituting "down" and combinations with "down."

The patterns were presented to *S*s on a memory drum. The *S*s viewed the words of the patterns through a slot in a large fiberboard screen separating *S* and *E*. The correct response was presented to *S* for 2 sec., but no fixed response time interval was required under either procedure. The lists were presented, in the manner indicated in the instructions given below, with a 30-sec. interval between presentations.

When *S* was seated in front of the screen, *E* read the instructions, handing *S* a small card on which the choices used in the pattern were printed. The words on the cards were pointed out to *S* as they occurred in the instructions. The *S*s of the 2-choice correction group were told they were to learn a series of words, the words being the directions "right" and "left." When signaled to start they were to name the directions, and when the correct direction was named it would appear on the memory drum in front of them. Directions were then to be named again until the second correct direction was named, whereupon it would appear, and so on throughout the list. The list was to be repeated until one errorless trial was completed. The instructions for the remaining correction groups differed only in terms of number and kinds of choices. For the modified correction groups, the instructions differed in that *S* was permitted to name only one direction, after which the correct direction would appear on the drum. The instructions were otherwise comparable.

One hundred *S*s, all volunteers from elementary classes in psychology at the University of Wisconsin, served in the experiment. None of them had had previous experience in this situation. Of the 100, 64 were women and 36 were men. The *S*s were assigned at random to each cell of the 2×5 factorial design employed. All *S*s learned to a criterion of one perfect trial.

Results

There are three possible scores, trials, errors, or time to the criterion, to employ in measuring acquisition. The error scores may be of two types: total errors and total first errors. The total first errors measure is used to compare the two procedures on the number of first errors per number of stimulus presentations. It includes all errors made under the noncorrection procedure, but only a part of the errors made under the correction procedure. For the purpose of comparing the two procedures, all these scores may be considered appropriate measures.

Satisfactory comparisons of acquisition rates for the various mazes, in

correction learning, may be made, as Brogden and Schmidt have pointed out (2), only in terms of time or error scores. Thus trials to the criterion are inadequate since successive trials are unequal in time and in terms of total number and kind of errors.

Analyses of variance were completed on the data for all four measures and the results are given in Table 1. It will be noted that the results for time and total errors are similar to each other, and those for trials and total first errors, while differing from time and total errors, are similar to each other. Examination of the means (see Fig. 1) shows that, in general, maze difficulty increases with number of alternate choices. For time and total errors, the correction procedure is more difficult than the modified correction procedure, but the two procedures are not significantly different in terms of trials and total first errors.

To compare acquisition rates under the two procedures, Vincent curves were constructed for the first-errors measure. The curves for the two procedures are essentially superimposable, indicating no differential effect of procedure on learning rate.

Both time and total errors to the criterion as a function of number of choices per choice point are plotted for each of the two procedures in Fig. 1. The orthogonal polynomials analysis (3) was completed on these two measures for each procedure, and in all four cases the linear component, and only the linear component, is significant. Thus it is evident that all the functions are best described by best-fitting straight lines. For the time data the equation of the line fitted to the correction group is

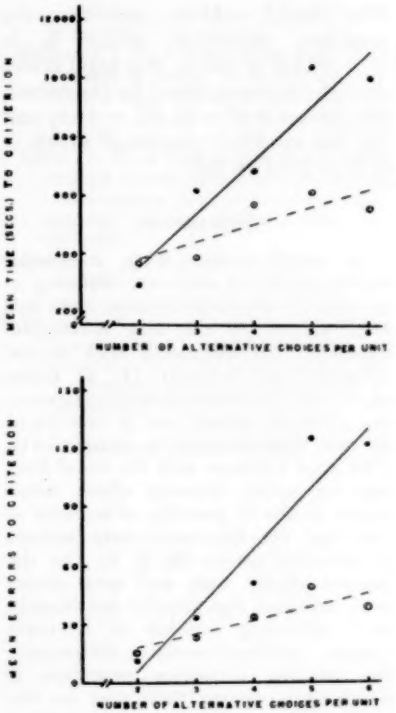
MEAN TIME (SECS.) TO CRITERION
NUMBER OF ALTERNATIVE CHOICES PER UNIT
MEAN ERRORS TO CRITERION
NUMBER OF ALTERNATIVE CHOICES PER UNIT

FIG. 1. Mean time and mean total errors to criterion as a function of number of alternate choices per maze unit. Solid dots and lines represent empirical points and best-fitting straight lines for the correction procedure; hollow dots and dashed lines represent empirical points and best-fitting straight lines for the noncorrection procedure.

TABLE 1

SUMMARY OF *F*-RATIOS FROM ANALYSES OF VARIANCE

Source	*df*	Measures			
		Time	Trials	Total Errors	Total First Errors
Mazes (M)	4	14.05**	5.34**	14.03**	14.26**
Procedures (P)	1	16.38**	0.34	22.59**	0.97
M × P	4	2.94*	1.21	5.47**	0.55

* $P = .05$.
** $P = .01$.

$Y = 186.6X - 21.9$, and for the modified correction group it is $Y = 59.8X + 261.5$. For total errors data the equation fitted for the correction group is $Y = 31.1X - 55.8$, and for the modified correction group it is $Y = 7.1X + 3.7$.

Discussion

As noted earlier, when comparing verbal mazes of this type differing in number of alternate choices, time and total errors are the only meaningful measures. In comparing such mazes, Brogden and Schmidt (**1, 2**) found significant increases with increasing number of choice points, and in fact found the data best described by straight lines. This is at variance with the usual findings on verbal learning where, when length of list or quantity of material is increased, the disproportionate increase in difficulty occurs (**4, 5, 6**). In the present study, time and total errors both increased significantly and linearly with increasing number of alternate choices. For these measures, the straight line for the correction procedure is significantly steeper than that for the noncorrection procedure, indicating much greater increase in difficulty with increase in the number of choices per choice point, as was predicted.

The results of the present study clearly demonstrate the superiority of the modified correction procedure over the correction procedure in terms of time and total errors to criterion. Trials are not comparable for the two procedures because of greater time and greater number of responses (and so, total errors) made in the correction procedure. Total first errors are even less comparable, since they include all errors made in the modified correction procedure but only a part of the errors made in the correction procedure. Thus these two measures *cannot* be considered satisfactory indices of learning per se for comparison of the two procedures. They do, however, provide indices concerning the number of presentations of the correct responses required by *S* to learn. Thus the number of presentations of the correct stimuli is twelve times the number of trials. Total first errors indicates how many times *S* requires correction—i.e., is given the correct stimulus after having made an error—to learn.

The discovery element, present in the correction procedure, is essentially eliminated in the modified correction procedure. Thus as number of alternate choices increases, increasing amount of discovery is required in the correction procedure while no increase is required in the modified correction procedure. In the former, greater time is required and many more responses, interpolated between presentations of the correct choices, are made. When such conditions apply in verbal learning, difficulty is increased (**4**). It is therefore surprising to note that, although time and total errors are indeed greater for correction learning, no more presentations of the correct responses are required by *S* than in modified correction learning. This would seem to argue that the discovery element has no significant effect upon the number of presentations of the correct responses required for learning, but serves rather to increase time and errors in proportion to the amount of discovery necessary. The fact that the Vincent curves for first errors are the same for both procedures demonstrates that the greater amount of discovery required in the correction procedure has no effect on rate of learning.

The experimental findings of this paper have been utilized by Simon (**7**) in the construction of a learning model. His conclusions, based on this and other studies, are in essential agreement with those stated above.

Summary

The effects of two procedures, correction and modified correction, were compared on the acquisition of a 12-unit verbal maze. The two procedures did not differ in terms of trials and total first errors, but the correction procedure, which involved more discovery, required greater time and total errors. Time and total errors

increased linearly with number of alternate choices for both procedures, the correction procedure giving a line of steeper slope. Rate of learning, as measured by first errors, was the same for both procedures.

References

1. Brogden, W. J., & Schmidt, R. E. Acquisition of a 24-unit verbal maze as a function of number of alternate choices per unit. *J. exp. Psychol.*, 1954, **48**, 335–338.
2. Brogden, W. J., & Schmidt, R. E. Effect of number of choices per unit of a verbal maze on learning and serial position errors. *J. exp. Psychol.*, 1954, **47**, 235–240.
3. Grant, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.
4. Hovland, C. I. Human learning and retention. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 613–689.
5. McGeoch, J. A., & Oberschelp, V. J. The influence of length of problem and of transfer upon rational learning and its retention. *J. gen. Psychol.*, 1930, **4**, 154–168.
6. Scott, T. C., & Henninger, L. L. The relation between length and difficulty in motor learning: A comparison with verbal learning. *J. exp. Psychol.*, 1933, **16**, 657–678.
7. Simon, H. A. Amounts of fixation and discovery in maze learning behavior. *Psychometrika*, 1957, **22**, 261–268.
8. Thompson, R. F. A comparison of correction and non-correction procedures on the acquisition of a 12 unit verbal maze. Unpublished master's thesis, Univer. of Wisconsin, 1953.

(Received January 13, 1958)

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

ALLEY LENGTH AND TIME OF FOOD DEPRIVATION IN INSTRUMENTAL REWARD LEARNING[1]

NORMA FREDENBURG BESCH

Columbus Psychiatric Institute and Hospital, Ohio State University

AND

WILLIAM F. REYNOLDS

Western Michigan University

One type of delay of reinforcement, in instrumental reward learning, which Hull discusses (4) is the within-chain case of delay. It involves response chains of different lengths or different distances between the initial portion of the chain and the reinforcing event at the end of the chain.

This variable of distance was investigated by Hull (2) in a situation in which he trained two groups of rats to run to food in a straight alley maze, one group in a 20-ft. maze and the other group in a 40-ft. maze. When running time was measured in successive 5-ft. sections, it was found that the running time in the initial section for the short alley group was less than that in the comparable section for the long alley group.

A factorial design was employed in a more recent study by Fredenburg (1) in which time of deprivation and alley length were varied. The data for Trial 1 and Trials 2, 3, and 4 of each day were treated separately due to pronounced warm-up effects. Significant differences in running speed, measured in a 1-ft. segment in the initial portion of the maze, were obtained in both comparisons between animals maintained on 3-hr. food deprivation and animals maintained on 22-hr. deprivation. The 2.75 ratio of long to short path did not produce a significant effect on running speeds; only Trial 1 of each day showed any effect of alley length ($P < .09$). While the difference between the running speeds for the long and short paths tended to be larger for the low-drive group than for the high-drive group, the interaction was not significant.

The interaction between time of deprivation and alley length has implications for alternative interpretations of within-chain delay of reinforcement. If a nonsignificant interaction is obtained, Hull's (3) habit interpretation of within-chain delay will be rejected in favor of Spence's (6) *K*-factor interpretation. The present study was designed to investigate this interaction by using conditions similar to those in the Fredenburg study, but with a larger long to short path ratio.

METHOD

Subjects.—The *Ss* were 80 experimentally naive female hooded rats from the colony maintained by the Department of Psychology at the State University of Iowa. Their ages ranged from 90–110 days at the beginning of the experiment. Twenty *Ss* were assigned randomly to each of four experimental groups.

Apparatus.—The apparatus consisted of a starting section, 1.25 ft. long, separated by two manually operated doors from a 16-ft. straight alley. The alley height and width were 3 in. and 6 in., respectively. A removable block with

[1] The research was conducted at the State University of Iowa under the direction of K. W. Spence.

a brass dish, .25 in. from the floor, was used as the goal piece. The block was inserted in the alley either 4 or 16 ft. from the starting box and 1 ft. from a retrace door.

The entire apparatus was painted a flat black with the exception of the floor, which was covered with rubber matting. The top of the maze was covered with glass and the goal sections with hinged pieces of Plexiglas. Wooden frames, fitted over the outside of the maze to eliminate grooves in the sides of the alley proper, served as the frames for the retrace doors at the entrance to both goal sections. A black curtain extended across the alley .5 in. in front of the first retrace door. All doors were operated from behind the start box by means of a pulley system made possible by wooden crosspieces above the maze.

The lighting arrangement was designed to assure dim, diffuse, and homogeneous lighting for the entire length of the maze. For illumination, four 40-w. bulbs were equally spaced 5.5 ft. above the maze. Suspended over the length of the maze, 1 ft. below the bulbs, was a 17 × 2-ft. piece of cardbdard whose sides were each bent up 6 in. A black curtain was draped over a wire frame 2 ft. above the maze.

Model 330 of the Photo Contact Relay, Hunter Mfg. Co., Inc., was used in both the starting time and running time circuits. The light which activated the units was provided by three adequately shielded 1.5-w. lamps mounted on the outside of the alley directly opposite the units. The light passed through round holes, .5 in. in diameter. Raising the second door in front of the starting section started a Standard Electric clock which was stopped when *S* interrupted the light beam to a photo cell 3 in. distant. Similarly, in the running time circuit, successive interruption of the light beams to two photo cells, which marked off a 1-ft. segment 1 ft. from the starting section, registered on another Standard Electric clock.

All components of the circuits which were superthreshold in noise were placed in an adjoining room. The clocks, enclosed in a semisoundproof box, were the only exception. A fan, placed on the floor behind the apparatus, provided a general masking noise.

Procedure.—The *S*s were given 15 days habituation to a 24-hr. feeding rhythm, during which each *S* was fed for 1 hr. a day in an individual feeding cage. The 22-hr. and 3-hr. deprived groups were fed at a time appropriate to the subsequent experimental runs.

On Day 1 of the experimental treatment, *S*s were introduced into the starting box. As soon as *S* faced in the direction of the runway, the first door was raised. Approximately .5 sec. later the second door was raised. Half of the *S*s received two pellets of Purina Laboratory Chow, .1 gm. each, at a goal 4 ft. distant, while the other half received the same reward at a goal 16 ft. distant. Within these two groups, half of the *S*s ran under a low hunger drive, based on 3 hr. of food deprivation, while the other half ran at a high-drive level based on 22 hr. of deprivation. The groups are designated as follows: Group 1, short alley—22 hr. of food deprivation (Sh-22); Group 2, long alley—22 hr. of food deprivation (Lg-22); Group 3, short alley—3 hr. of food deprivation (Sh-3); and Group 4, long alley—3 hr. of food deprivation (Lg-3).

Two trials a day, at least 10 min. apart, were given to each *S* for 27 days. On Days 25–27, all *S*s were prefed in the goal box prior to the first experimental run. The order of runs for the groups was assigned randomly at the beginning of the experiment but remained constant thereafter. Clock readings were taken after each trial to the nearest .01 sec. At the end of each run, upon consumption of the reward, the animal was removed to an individual compartment in a carrying cage.

TABLE 1

MEANS AND *SD*'S OF SPEEDS ON TRIALS 1 AND 2 OF EACH DAY FOR TRIALS 1–48

Groups	Running Speed				Starting Speed			
	Trial 1		Trial 2		Trial 1		Trial 2	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
1 (Sh-22)	2.16	.76	2.66	.69	2.47	.68	2.59	.70
2 (Lg-22)	1.55	.70	2.26	.72	1.78	.72	1.93	.63
3 (Sh-3)	1.65	.63	2.24	.64	1.98	.49	2.17	.64
4 (Lg-3)	1.49	.74	2.38	.82	1.59	.51	1.64	.53

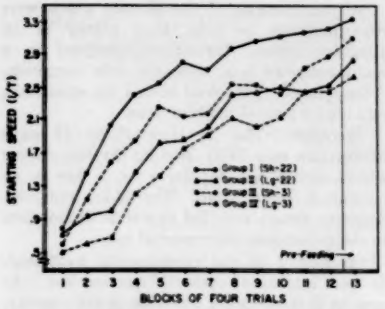

FIG. 1. Mean starting speeds in four trial blocks for Trials 1–52.

RESULTS

The means and *SD*'s of the running and starting speeds for all groups on Trials 1 and 2 of each day for Trials 1–48 are presented in Table 1. While both starting and running speeds were greater on Trial 2 of each day than on Trial 1, for starting speeds the increases were not a function of either alley length or time of deprivation. For the running speed measures, as in the Fredenburg study (1), alley length differences disappeared on Trial 2 of each day. In view of the consistent results over trials obtained with the starting speed measures, these data were selected for presentation in Fig. 1.

It is clear from an inspection of Fig. 1 that both alley length and deprivation time affected performance levels. The analysis of variance (5), presented in Table 2, yielded *F*'s of 18.22 ($P < .001$) and 6.77 ($P < .025$) for alley length and deprivation time, respectively. Since the main effects of the experimental treatments did not differ with respect to Trials 1 and 2 of each day, the trials variable was ignored in the *t* tests of the differences between groups. All comparisons yielded *t* values significant at the .05 level with the exception of the comparison between the long alley groups (Groups 2 and 4).

The analysis of variance of the running speed data, presented in Table 3, failed to yield significant *F*'s for either of the experimental conditions. However, *t* tests, justified by the significant trial by alley length interaction ($P < .005$) indicate that the alley length factor had a significant effect on Trial 1 of each day ($P < .025$).

TABLE 2

ANALYSIS OF VARIANCE OF MEAN STARTING SPEEDS FOR TRIALS 1–48

Source	*df*	*MS*	*F*
Total between *Ss*	79		
Deprivation (B)	1	2,775.19	6.77*
Alley length (C)	1	7,466.34	18.22**
B × C	1	247.76	
Error	76	409.80	
Total within *Ss*	80		
Trials* (A)	1	380.40	12.27**
A × B	1	2.37	
A × C	1	15.98	
A × B × C	1	42.22	
Residual	76	31.01	

* The trials variable is based on Trials 1 and 2 of each day.
* $P < .025$.
** $P < .001$.

TABLE 3

ANALYSIS OF VARIANCE OF MEAN RUNNING SPEEDS FOR TRIALS 1–48

Source	*df*	*MS*	*F*
Total between *Ss*	79		
Deprivation (B)	1	1,085.59	1.97
Alley length (C)	1	1,496.48	2.72
B × C	1	1,412.92	2.57
Error	76	550.43	
Total within *Ss*	80		
Trials (A)	1	10,393.94	270.37**
A × B	1	115.06	2.99
A × C	1	373.45	9.71*
A × B × C	1	14.16	
Residual	76	38.44	

* $P < .005$.
** $P < .001$.

Neither the starting speed nor the running speed measures yielded significant interactions between deprivation time and alley length. To facilitate consideration of these interactions, they are presented graphically in Fig. 2 and 3. Consistent with the nonsignificant interaction found for the starting speeds, the lines connecting the drive levels for the two alley length conditions are parallel. While an interaction appears to exist between drive level and alley length in the case of the running speed measures, particularly for Trial 2 of each day, it did not reach statistical significance ($.05 < P < .10$). It will be recalled that in the previous study the larger difference between the alley length groups favored the low-drive condition, which is opposite to the result found here.

If the feeding in the goal box after the first daily run of the Learning Period was responsible for the increased speeds observed on the second daily trial, a prefeeding procedure would be expected to produce the same effect. Analyses of variance were performed which compared the Prefeeding Period with the last three days of the Learning Period. Both running and starting speeds on the first trials of each day were significantly faster ($P < .001$) for the Prefeeding Period. The differences between the running speeds on the second trials of the Learning Period and the first trials of the Prefeeding Period were not significant ($F = 1.52$, $P > .20$). However, for starting speeds, the same comparisons yielded a significant F ($F = 11.23$, $P < .005$); that is, starting speeds were greater for the Prefeeding Period.

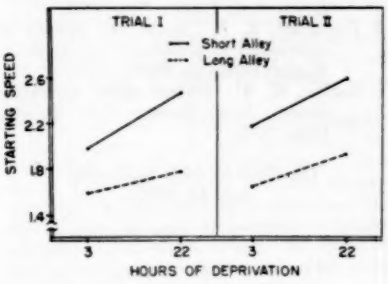

FIG. 2. Mean starting speeds for Trials 1–48.

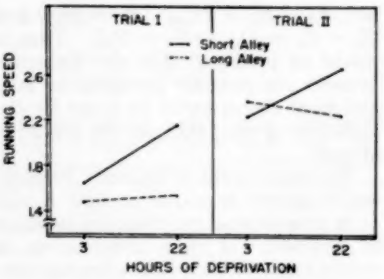

FIG. 3. Mean running speeds for Trials 1–48.

DISCUSSION

While Hull (3) attributes the difference in response strength for long and short chains to a difference in habit strength at the beginning of each chain, Spence (6) interprets within-chain delay of reinforcement in terms of the incentive motivating factor, K. From the Hullian position, in which it is assumed that (*a*) the habit strength of the initial S-R component of a response chain is a function of the length of the chain, (*b*) differential drive is a function of food privation, and (*c*) drive level and habit strength combine multiplicatively, the excitatory potentials for the short (E_S) and long (E_L) alleys would be as follows:

$$E_S = H_S \times D$$
$$E_L = H_L \times D$$

The differences between the excitatory potentials for the two alley lengths for the low-drive groups and the high-drive groups, respectively, would then

be: $E_S - E_L = D_{Low}\ (H_S - H_L)$ and $E_S - E_L = D_{High}\ (H_S - H_L)$. Thus, it would be predicted that the difference between the response strengths for long and short chains would be larger for the high-drive groups than for the low-drive groups.

The implications of Spence's interpretation depend primarily on how K and D, as motivational variables, are assumed to combine. If they combine in a simple linear fashion the following relationships are obtained:

$$E_S = H\,(D + K_S)$$
$$E_L = H\,(D + K_L)$$
$$E_S - E_L = H\,(K_S - K_L)$$

Thus, when it is assumed that habit strength combines multiplicatively with total drive which is an addition of all drives present, the resulting response strengths are a function only of the values of K, or are independent of hunger drive. One would expect, then, that the differences between the long and short alley groups would be the same for both drive conditions.

The experimental evidence appears to be more consistent with Spence's position. It will be recalled that the interactions between food deprivation and alley length were nonsignificant in all comparisons. However, the "warm-up" effect, the increased speeds on the second daily trials of the Learning Period, is problematic for either interpretation. From the comparisons between the Prefeeding Period and the Learning Period, it is suggested that the effect was due to the first trial feedings. While goal-box behavior is critical in Spence's treatment of the K-factor, such rapid changes in performance would not be expected. Also, there is no a priori reason to expect the changes to differ with length of delay, as was found for running speeds. Similarly, habit is conceived by Hull to develop gradually as a function of the number of reinforcements rather than being subject to rapid shifts in its strength.

The problem remains that there was a lack of consistency in the present experiment between running and starting speed measures. However, it might be argued that starting speed was a more sensitive measure, since it yielded significant and consistent effects for both experimental variables.

Summary

Eighty rats were trained to run to food in a straight alley. A factorial design was employed with two lengths of alley and two levels of food deprivation. Running speeds were significantly faster for the short alley than for the long alley groups on Trial 1 of each day only, while for starting speeds the same effect was significant on both daily trials. Deprivation time had a significant effect only for starting speeds. The interaction between deprivation time and alley length was insignificant for both measures. The results were discussed in terms of Hull's habit interpretation and Spence's K-factor interpretation of within-chain delay of reinforcement.

References

1. Fredenburg, N. C. Response strength as a function of alley length and time of deprivation. *Proc. Iowa Acad. Sci.*, 1956, **63**, 563–567.
2. Hull, C. L. The rat's speed-of-locomotion gradient in the approach to food. *J. comp. Psychol.*, 1934, **17**, 393–422.
3. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
4. Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
5. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
6. Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

(Received for early publication June 10, 1958)

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

CORRECTION AND REANALYSIS

NORMAN H. ANDERSON[1] AND DAVID A. GRANT
Yale University *University of Wisconsin*

W. K. Estes has kindly called our attention to an error in the Anderson-Grant article (**2**) where it was stated that the quantity b was an estimate of $\theta(2\alpha_j - 1)$. The derivation of this result failed to take account of the information about the underlying response probability which is conferred by knowledge of the last response made. This necessitates a reconsideration of two conclusions of the earlier report: (*a*) that the quantities a and b were inconsistent as estimates of θ; and (*b*) that b could be interpreted as a statistic descriptive of a perseverative tendency.

Fortunately, the situation can be handled correctly with the general estimation method for linear learning models developed by the first author (**1**) which uses the sequential dependencies within the sequences of stimuli and responses. In the context of the model in the earlier article, the relevant equations take the following special form:

$$p_{i,11} = \pi_i[\theta\alpha_j V_1 + (1 - \theta)V_2]$$

$$p_{i,21} = \pi_i[\theta(1 - \alpha_j) + (1 - 2\theta + \theta\alpha_j)V_1 - (1 - \theta)V_2]$$

$$p_{i,12} = \pi_i[(1 - \theta\alpha_j)V_1 - (1 - \theta)V_2]$$

$$p_{i,22} = \pi_i[1 - p_{i,11} - p_{i,21} - p_{i,12}]$$

Here π_i is the probability of the noncontingent reinforcing event, E_i; θ is the learning rate; α_j is the sharing parameter; V_1 and V_2 are the mean and second raw moments of the distribution of response probability, averaged over trials; $p_{i,21}$ is the *unconditional* probability of a subsequence of the form (Response A_2 on Trial n, Event E_j on Trial n, Response A_1 on Trial $n + 1$), with analogous definitions for the other p's. The p's are to be considered as averaged over the population of organisms on any given trial, and then averaged over trials. It may be noted that the same equations hold for single event trials (by setting $\alpha_j = 1$), and for event trials on which no learning occurs (taking $\theta = 0$).

Lines 1, 2, and 3 of Table 1 give the observed values of $p_{i,11}/\pi_i$, $p_{i,21}/\pi_i$, and $(p_{i,21} - p_{i,12})/\pi_i$, respectively, for the double event trials, $j = 0$ or 3. These values were obtained directly from the 2×2 tables of the frequencies of the four possible doublets of successive responses separated by a double event. All trials on which a double event occurred were used in the tabulations.

Present interest centers on the differences in Lines 3 where a positive value means that the frequency of A_1 responses is higher following double events than it is preceding double events. The analysis of these differences, performed on the raw frequencies, is in Table 2. For Groups .30 and .45, the differences are uniformly positive and are significantly different from zero as shown by the F's for Mean. The decreasing trend over trials is significant only for Groups .30. For Groups .60, the mean differences are small and nonsignificant. There is no indication that type of double event has an effect.

Since V_1 is less than .5 in each trial block for Groups .30 and .45, (**2**, Fig. 4), the analysis demonstrates a significant increase in the probability of the less frequent A_1 response following double-event trials. Since V_1 is not much different from .5 for Groups .60, and since this effect does not appear here, it would evidently be better interpreted as a regression toward the chance level rather than as a regression toward the A_1 response which was conditioned on Trials 1–100.

TABLE 1

OBSERVED AND PREDICTED SEQUENTIAL DEPENDENCIES

Group	Line	Trial Blocks		
		101–200	201–300	301–400
.30-0	1	.192	.165	.109
	2	.162	.125	.111
	3	.088	.068	.035
	4	.020	.019	.016
.30-3	1	.217	.133	.099
	2	.160	.126	.119
	3	.088	.049	.035
	4	.015	.017	.016
.45-0	1	.314	.239	.203
	2	.193	.164	.134
	3	.051	.043	.023
	4	.006	.011	.014
.45-3	1	.268	.221	.178
	2	.165	.142	.116
	3	.031	.028	.019
	4	.011	.011	.011
.60-0	1	.405	.317	.364
	2	.114	.134	.138
	3	−.003	−.006	.007
	4	.000	.003	.001
.60-3	1	.361	.330	.373
	2	.137	.133	.147
	3	.008	.006	.021
	4	.000	.003	.000

Note.—See text for explanation of entries.

[1] Now at the University of California, Los Angeles.

TABLE 2

ANALYSIS OF VARIANCE OF SEQUENTIAL DEPENDENCY DIFFERENCES

Source	*df*	*F*'s for Double-Event Frequency Groups		
		.30	.45	.60
Between				
Mean	1	27.24*	17.54*	1.41
Groups	1	0.10	0.67	1.65
Error	46	(16.97)	(17.29)	(11.37)
Within				
Trials	2	4.72*	1.04	0.97
T × G	2	0.18	0.15	0.01
Error	92	(6.73)	(10.73)	(9.58)

Note.—Error mean squares in parentheses.
* $P < .05$.

The quantity b is obtainable as a conditional probability from Table 1. However, since b involves an estimated denominator, it is preferable to work directly with the unconditional probabilities. The question as to the consistency of a and b is then equivalent to asking whether θ as estimated by a will satisfy the equations above. The simplest check is gotten by subtracting the third equation from the second, and dividing by π_j to obtain $\theta(1 - \alpha_j)(1 - 2V_1)$ as the theoretical expression for the differences of Lines 3 of Table 2. The predicted regression thus depends on the learning rate, the sharing parameter, and the distance from the chance level of response.

This expression may be evaluated from the earlier data (2). The sharing parameters were assigned the common value, .88, since type of double event had no effect in the above analysis. Using the obtained values for V_1, and substituting a for θ, (2, Fig. 4 and 5), this expression then yields the predicted values in Lines 4 of Table 2. It is seen that they are rather definitely too small compared to the observed values in Lines 3. The inconsistency of a and b thus remains unresolved.

The validity of interpreting b as descriptive of a perseverative tendency is unclear. In general, it is evident that interpretations of this kind must depend on some assumed model, explicit or implicit, which provides a null-hypothesis baseline. (The baseline used in the earlier article was the mean proportion of A_1 responses, which is invalid as soon as it is admitted that response probability changes or fluctuates from trial to trial.) This unfortunately makes the intuitive meaning of b and analogous statistics model-dependent, but there seems to be no way to avoid the problem. Lacking an adequate model for the situation, it is necessary to view the earlier interpretation of b circumspectly.

The regression effect discussed above is a source of difficulty to all the models which have been proposed for this experimental situation. The present sharing parameter model predicts values which are considerably too small. As presently formulated, the models of Neimark (4) and Atkinson (3) predict zero effect. It would seem that this discrepancy cannot be attributed to nonindependence of successive samples since the regression does not appear for Groups .60. It may be expected that a more extended investigation of the sequential dependencies will shed additional light on the problem.

REFERENCES

1. ANDERSON, N. H. An estimation method for linear learning models. Paper presented at the Social Science Research Council Summer Institute, Stanford Univer., 1957.
2. ANDERSON, N. H., & GRANT, D. A. A test of a statistical learning theory model for two-choice behavior with double stimulus events. *J. exp. Psychol.*, 1957, **54**, 305–317.
3. ATKINSON, R. C. An analysis of the effect of non-reinforced trials in terms of statistical learning theory. *J. exp. Psychol.*, 1956, **52**, 28–32.
4. NEIMARK, E. D. Effects of type of nonreinforcement and number of alternative responses in two verbal conditioning situations. *J. exp. Psychol.*, 1956, **52**, 209–220.

(Received for early publication July 21, 1958)

ERRATUM

In the article "Dimensional Analysis of Movement Reactions" by Edwin A. Fleishman, in the May, 1958, issue of this JOURNAL (*J. Exp. Psychol.*, **55**, 438–453), several reference numbers in the text should be changed to conform with the References at the end. On p. 439, 5th line from bottom, second column, Reference **7** should be Reference **10**. On p. 442, the correct reference for Test 2 is Reference **13** instead of **12**. On p. 443, the references for Tests 7, 8, and 10 should be Reference **13**.

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

SUPPLEMENTARY REPORT: PROCESSES UNDERLYING LEARNING A SINGLE PAIRED-ASSOCIATE ITEM[1]

JOHN OLIVER COOK

North Carolina State College

As a test of the theoretical model of paired-associate learning previously outlined (1), a second paired-associate experiment was run on 160 airmen, with eight groups of 10 in each treatment. As in the previous experiment, the same experimental materials were used, and the same two learning methods: Prompting (stimulus term, then response term, then overt practice of the response term) and Confirmation (stimulus term, overt practice of the response term, and then the response term). The superiority of the Prompting method, and hence of the model that predicts its superiority, had been questioned on the grounds that the highest level of learning achieved in the previous experiment was only 54.8% of the material. Had training been continued, it is alleged, Confirmation might have proved superior in the later stages.

To meet this objection *Ss* in the present experiment were run to mastery with a test after every three of the 36 training trials. In terms of the number of correct responses, as the theoretical model predicted, Prompting proved superior throughout the learning curve (Table 1). Prompting also proved superior in terms of the number of different "legitimate" responses (a "legitimate" response being any one of the 10 response terms, whether elicited by its proper stimulus term or by one of the others). In the previous experiment this finding appeared only as a tendency but did not achieve significance. Prompting produced more substitution errors early in learning because *Ss* knew a larger number of response terms ("legitimate" responses), but the number of substitution errors, as the model predicted, dropped off rapidly. In the Confirmation condition, substitution errors reached their peak later and showed a significant tendency to persist, due to the fact that the Confirmation method makes it difficult for the implicit response to the response term to become attached to the stimulus consequences of the implicit response to the stimulus term.

[1] This report is based on work done under ARDC Project No. 7709, Task No. 37305, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas.

REFERENCE

1. COOK, J. O., & KENDLER, T. S. A theoretical model to explain some paired-associate learning data. In G. Finch and F. Cameron (Eds.), *Symposium on Air Force human engineering, personnel, and training research.* Washington, D. C.: National Academy of Science—National Research Council, Publ. No. 455, 1956. Pp. 90–98.

(Received January 13, 1958)

TABLE 1

NUMBER OF CORRECT RESPONSES, DIFFERENT "LEGITIMATE" RESPONSES, AND SUBSTITUTION ERRORS
(Means of medians of eight groups in each of two treatments)

Test	Correct Responses			Different Legitimate Responses			Substitution Errors		
	Prompt.	Confirm.	*P**	Prompt.	Confirm.	*P**	Prompt.	Confirm.	*P**
1	.250	.000	—	3.000	2.125	.02	3.250	2.375	.05
2	1.750	1.000	.01	4.250	2.875	.01	3.000	2.250	.05
3	3.625	2.375	.01	5.875	4.000	.01	1.750	2.125	—
4	5.875	3.875	.02	6.625	5.000	.01	.625	1.500	.02
5	7.625	5.500	.02	8.000	6.375	.05	.125	.875	.01
6	9.000	6.750	.01	9.000	7.375	.02	0	.625	.05
7	9.625	8.000	.02	9.625	8.250	.05	.125	.375	—
8	10.000	9.000	.05	10.000	9.250	.05	0	0	—
9	10.000	9.750	—	10.000	9.750	—	0	0	—
10	10.000	10.000	—	10.000	10.000	—	0	0	—
11	10.000	10.000	—	10.000	10.000	—	0	0	—
12	10.000	10.000	—	10.000	10.000	—	0	0	—

* *P* values determined by Kruskal-Wallis rank test, using medians of the eight Prompting groups and medians of the eight Confirmation groups.

Journal of Experimental Psychology
Vol. 56, No. 5, 1958

SUPPLEMENTARY REPORT: THE DIMINISHING INTERVAL CONDITIONING TECHNIQUE

ROBERT E. MILLER AND JOHN V. MURPHY

School of Medicine, University of Pittsburgh

The technique of progressively diminishing the intertrial interval in avoidance conditioning has not only the practical advantage of considerable savings in testing time but also, of even greater significance, the avoidance response established under this procedure has been shown to be significantly more stable and resistant to extinction than a response established under the conventional nondiminishing procedure (1). The present study was designed to study the course of extinction of this remarkably stable response over an extended period. In addition to a group conditioned for 150 trials on a diminishing schedule, as in the previous experiment, a second group conditioned for only 100 trials was included.

METHOD

Subjects.—Thirty male rats of the Carworth strain were used. The *S*s were 50–60 days old at the beginning of training.

Apparatus.—The apparatus consisted of a two-compartment avoidance conditioning box (1). The two compartments were separated by a 2-in. barrier.

Procedure.—An avoidance response, barrier crossing, was conditioned to a buzzer CS in accord with the method of the previous report (1). Ten conditioning trials were presented each day. The average intertrial interval was diminished in 10-sec. steps from 60 sec. to 20 sec. during conditioning. The *S*s were randomly assigned to either a 10- or 15-day conditioning group. The average interval was diminished 10 sec. every third day for the 10-day group and every fourth day for the 15-day group.

Following conditioning, *S*s were tested at the rate of 10 trials per day for a 72-day period on an extinction procedure, i.e., no shock was administered for failure to respond. The intertrial interval during extinction averaged 20 sec. for each group.

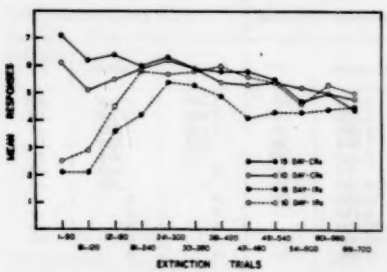

FIG. 1. The incidence of conditioned avoidance responses and interval responses during extinction.

RESULTS AND DISCUSSION

Three *S*s in the 10-day group and two *S*s in the 15-day group were discarded prior to extinction testing for failure to meet the criterion of conditioning (1). The mean incidence of avoidance responses for the remaining *S*s in successive blocks of 60 extinction trials is shown in Fig. 1. A repeated measures analysis indicated that there was no significant difference during extinction between the 10-day and the 15-day groups. There was, however, a downward trend in the number of avoidance responses during the extinction period ($P < .05$).

These data indicate that the conditioned avoidance response established with the diminishing interval technique had not approached extinction despite the extended period of testing. It appears that many hundreds of additional trials would be required to extinguish this response to zero level. For all practical purposes the avoidance conditioning established with this procedure might well be considered nonextinguishable.

Mean interval responses during the 72-day extinction period are shown in Fig. 1. These responses are, as the name implies, barrier-crossings which occur between trials and in the absence of either CS or UCS. A repeated measures analysis showed no significant difference between the 10-day and 15-day groups. The incidence of interval responses increased significantly ($P < .01$) during the extinction phase.

REFERENCE

1. MURPHY, J. V., & MILLER, R. E. Spaced and massed practice with a methodological consideration of avoidance conditioning. *J. exp. Psychol.*, 1956, **52**, 77–81.

(Received February 4, 1958)

The uses of apparatus and laboratory procedure in psychological experiments.

Laboratory Instrumentation in Psychology

by W. W. Grings
University of Southern California

An introduction both for the student and for the professional man or woman in behavior research who seeks detailed information and evaluation on the use of equipment and techniques.

282 pages, 142 illustrations, cloth, $4.75

Order from your bookstore or from

N-P PUBLICATIONS, The National Press, 850 Hansen Way
Palo Alto, California

A *very considerable* proportion of all published articles on perceptual-motor skills and on perception-as-influenced-by-experience are given immediate publication in

PERCEPTUAL AND MOTOR SKILLS

If you are interested in these fields, you and your library should have current subscriptions to PMS ($10.00 per year, with a special rate of $5.00 to individual members of the APA). Prices of back volumes may be obtained on request from

PERCEPTUAL AND MOTOR SKILLS
PSYCHOLOGICAL REPORTS
Box 1441
Missoula, Montana